Imperial College London

The Division of Psychiatry, Department of Medicine

The Development, Assessment and Treatment of Behaviour Problems in Young Children: Integrating Children's Own Perspectives

Bethany Barker

Supervisors: Professor Paul Ramchandani and Dr Christine O'Farrelly

Funded by the Imperial College London President's PhD Scholarship

Submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy

Word count: 85,178

July 2021

Statement of Originality

I declare this work to be my own. Any inclusion of the work or ideas of others is appropriately referenced. Others' contributions to the research process are acknowledged throughout.

B Barker

21st July 2021

Copyright declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Background

Studies of child development typically rely on the insights of adults to understand children's behaviour. The value of integrating children's own perspectives into these investigations has been more widely recognised in recent years. However, such opportunities are rarely extended to very young children, particularly in populations exhibiting challenging behaviours. This thesis contributes empirical evidence to this field by obtaining child-reported data in a large sample of three- and four-year-olds. Using a play-based story stem battery, children's narrative representations and their relation to children's behaviour, parental discipline, parental sensitivity, and a video-feedback parenting intervention are explored.

Setting, Participants, and Main Measures

This research was embedded in the *Healthy Start, Happy Start* study, a pragmatic randomised controlled trial. Participants were 300 children, who were 12-36 months old and exhibited elevated behaviour problems at screening, and their caregivers. Families were assessed at baseline, 5-month follow-up, and 24-month follow-up. In this thesis, key measures were child behaviour assessed at all time-points, parental sensitivity assessed at baseline and 5-month follow-up, and children's narrative representations assessed at 24-month follow-up.

Results

Results from three studies are presented. The first, a cross-sectional study, revealed modest but consistent associations between constructs derived from children's narratives and parent- and teacher-reported behaviour. The second and third studies were longitudinal and experimental in design. Behaviour problems in toddlerhood were associated with disruptive story themes at three- and four-years-old. Early parental sensitivity was associated with narrative coherence and observed behavioural dysregulation at three- and four-years-old. A video-feedback parenting intervention was found to have small to medium, positive effects on children's behaviour problems, narrative coherence, and expression of prosocial themes. There was no evidence that improvements to parental sensitivity mediated these relationships.

Conclusion

This thesis contributes to our understanding of the influences of early behaviour and parent-child interactions on children's development and highlights the feasibility and unique value of integrating young children's own perspectives into such investigations.

Table of contents

Table of contents **5**

List of tables **11**

List of figures **14**

List of appendices **15**

Acknowledgements **16**

1 CHAPTER 1: Behaviour problems, the parent-child relationship, and children's voices: An overview **18**

1.1 Chapter rationale 18

1.2 Introduction 19

1.3 Theoretical frameworks and models 21

1.3.1 Attachment theory 21

1.3.2 Social learning and coercion theories 21

1.3.3 Ecological models of development 22

1.3.4 Revisionist vs. enduring effects models of development 22

1.3.5 Rights-based and democratic frameworks of early childhood. 23

1.4 Behaviour problems 24

1.4.1 What are behaviour problems? 24

1.4.2 What is the prevalence of these difficulties? 25

1.4.3 What challenges can long-term difficulties present? 27

1.4.4 When to intervene? 28

1.4.5 How to intervene? 29

1.5 Parenting and the parent-child relationship 30

1.5.1 What components of parenting are important for children's behavioural development? 30

1.5.2 What are some of the limitations in our understanding of the parent-child relationship? 32

1.5.3 How can we measure parenting? 34

1.5.4 How can parenting interventions promote family functioning? 35

1.6 Integrating children's perspectives 40

1.6.1 Why and how should we listen to children? 41
1.6.2 What are some of the key ethical considerations when integrating children's perspectives? 44
1.6.3 What are some of the methodological challenges of integrating children's perspectives? 44
1.6.4 How can young children's perspectives be elicited? 45
1.7 The current thesis 48
2 CHAPTER 2: Healthy Start, Happy Start: An overview 49
2.1 Chapter rationale 49
2.2 Introduction 50
2.3 Method 52
2.3.1 Study design 52
2.3.2 Ethics and trial governance 52
2.3.3 Eligibility criteria 52
2.3.4 Participant recruitment 53
2.3.5 Sample characteristics 57
2.3.6 Informed consent 59
2.3.7 Baseline and follow-up assessments 60
2.3.8 Measures 60
2.3.9 Randomisation 66
2.3.10 Intervention arm: VIPP-SD 66
2.3.11 Control arm: Usual care 67
2.3.12 Blinding 67
2.3.13 Participant retention 67
2.3.14 Data monitoring and quality assurance 68
2.3.15 The use of Healthy Start, Happy Start data in this thesis 68
3 CHAPTER 3: Exploring the use of a child-informant measure of child behaviour in a sample of three- and four-year-olds 69
3.1 Chapter rationale 69
3.2 Introduction 70

3.2.1 Children as informants 70
3.2.2 Doll-play story stem narrative assessments 72
3.2.3 Conclusion 79
3.2.4 Present study research aims 79
3.3 Method 80
3.3.1 Design 80
3.3.2 Participants 80
3.3.3 Measures 80
3.3.4 Study procedure 92
3.3.5 Statistical analysis 92
3.4 Results 94
3.4.1 Description and exploration of study data 94
3.4.2 Aim 1: Examining the feasibility and acceptability of delivering the measure 100
3.4.3 Aim 2: Exploring concurrent associations between adult-reported child behaviour and constructs derived from children's responses to the story stems. 106
3.4.4 Aim 3: Explore whether child behaviour, measured through the story stem assessment, varied based on characteristics that typically predict child behaviour 109
3.5 Discussion 117
3.5.1 Summary of findings 117
3.5.2 Strengths and limitations 122
3.5.3 Implications and conclusions 127
4 CHAPTER 4: Exploring associations between early child behaviour, parent-child interactions, and children's later narrative representations 129
4.1 Chapter rationale 129
4.2 Introduction 130
4.2.1 Theoretical basis to exploring children's internal representations 131
4.2.2 Methods and approaches to accessing internal representations 133
4.2.3 Links between external and internal worlds 136
4.2.4 Conclusion and current study's research aims 139
4.3 Method 141

4.3.1 Study design 141
4.3.2 Participants 141
4.3.3 Measures 142
4.3.4 Study procedure 147
4.3.5 Intervention: VIPP-SD 147
4.3.6 Statistical analysis 148
4.4 Results 150
4.4.1 Preliminary analysis 150
4.4.2 Main analysis 160
4.4.3 Does early parental discipline and child behaviour predict children's later expression of disruptive story themes? 163
4.4.4 Does early parental discipline and child behaviour predict children's later observed behaviour dysregulation during the completion of a story stem task? 165
4.4.5 Does early parental discipline and child behaviour (prosociality and behaviour problems) predict children's later expression of prosocial story themes? 167
4.4.6 Does early parental discipline and child behaviour (prosociality and behaviour problems) predict children's later narrative coherence? 170
4.4.7 Considering the effects of a video-feedback parenting intervention on children's story stem outcomes 173
4.5 Discussion 174
4.5.1 Summary of findings 174
4.5.2 Strengths and limitations 179
4.5.3 Implications and conclusions 182
5 CHAPTER 5: The place of sensitivity: Examining associations between parental sensitivity, behaviour problems, and children's narrative representations in the context of a video-feedback parenting intervention 183
5.1 Chapter rationale 183
5.2 Introduction 184
5.2.1 Parenting and children's development 184
5.2.2 Sensitive caregiving 185
5.2.3 Parental sensitivity in the context of child outcomes and intervention 192

5.2.4 Measuring intervention effects through children's representations 198
5.2.5 Aims of the current study 201
5.3 Method 202
5.3.1 Study design 202
5.3.2 Participants 202
5.3.3 Measures 202
5.3.4 Study procedure 210
5.3.5 Intervention 211
5.3.6 Research aims and planned statistical analysis 212
5.4 Results 216
5.4.1 Data included in current analysis 216
5.4.2 Power analysis 216
5.4.3 Descriptive analysis of the study sample and study variables 217
5.4.4 Aim 1: To explore associations between early parental sensitivity and children's later narratives 227
5.4.5 Aim 2: To explore whether intervention effects on narrative coherence were mediated by parental sensitivity 237
5.4.6 Aim 3: To explore whether intervention effects on children's behaviour (as measured through the CBCL) were mediated by parental sensitivity. 245
5.5 Discussion 253
5.5.1 Overview and discussion of findings 253
5.5.2 Strengths and limitations of the study 260
5.5.3 Conclusion and implications 265
6 CHAPTER 6: Discussion 267
6.1 Chapter rationale 267
6.2 Chapter overview 268
6.3 Summary of findings 268
6.3.1 Including children's perspectives is feasible, meaningful, and valuable 270
6.3.2 The parent-child relationship is an important factor in children's behavioural development 271

6.3.3 Parenting supports delivered in children's earliest years have the potential to improve family functioning 272
6.3.4 Parenting is a complex construct and just one of many contributing factors to children's development 272
6.3.5 Aspects of the richness and nuance of family life and child development are absent in these findings 273
6.4 Reflections on the strengths and limitations of this thesis 274
6.4.1 Notable strengths 275
6.4.2 Notable limitations 278
6.5 Implications for research and practice 281
6.5.1 Very young children's perspectives can be included in large-scale research 281
6.5.2 Early intervention may offer a powerful way to improve children's internal and external functioning 282
6.5.3 A single intervention will not address the dynamic needs of families 283
6.5.4 Systemic and entrenched issues obstruct families' capacity to access, engage in, and benefit from support 284
6.6 Conclusion 286
References 288
Appendices 329

List of tables

Table 1 *Overview of Parenting Interventions Tested with Families of Young Children* 37
Table 2 *Methodological Approaches to Elicit the Perspectives of Young Children with Exemplar Research Cases* 46
Table 3 *Recruitment Pathway for Participants Recruited into the HSHS Trial* 56
Table 4 *Recruitment Location of Participating HSHS Families* 56
Table 5 *Sample Characteristics of Participating Children and their Primary Caregivers* 58
Table 6 *Timing of Data Collection for Variables Analysed in this Thesis* 65
Table 7 *Overview of Measures Developed to Explore Children's Narrative Representations* 75
Table 8 *Story Stem Scenarios and Prompts Presented to the Child in the Story Stem Measure* 84
Table 9 *Exemplar Descriptors for a Selection of Scores for Children's Narrative Coherence* 86
Table 10 *Demographic Information for Participating Children and their Primary Caregivers* 95
Table 11 *Descriptive Statistics for Study Variables Collected at the Two-Year Follow-Up* .. 98
Table 12 *Story Stem Battery Completion Status for Full HSHS Sample* 100
Table 13 *Illustrative Data Related to the Constructs Derived from the Story Stem Measure* 102
Table 14 *Internal Consistency of Story Stem Subscales* 105
Table 15 *Bivariate Correlations between Story Stem Variables and Parent-Reported Child Behaviour* 107
Table 16 *Bivariate Correlations between Story Stem Variables and Teacher-Reported Child Behaviour* 108
Table 17 *Bivariate Correlations between Story Stem Variables and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment* 111
Table 18 *Bivariate Correlations between Parent-Rated Child Behaviour and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment* 112
Table 19 *Parent- and Teacher-Reported Child Behaviour by Child Gender* 115
Table 20 *Bivariate Correlations between Child Age and Story Stem Constructs* 116
Table 21 *Sociodemographic Characteristics of the Full Sample and Split by Group Allocation* 152
Table 22 *Descriptive Statistics for Study Variables Included in the Current Analyses* 155

Table 23 *Bivariate Correlations between Measures of Child Behaviour and Parenting Practices Measured at Baseline and 24-Months* 157
Table 24 *Bivariate Correlations between Baseline Parent-Report Measures of Children's Behaviour and Parenting and Child Outcomes at 24-Month Follow-Up* 159
Table 25 *Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Disruptive Themes in Children's Stories Two Years Later* 164
Table 26 *Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Observed Behavioural Dysregulation Two Years Later* 166
Table 27 *Linear Model of Early Parental Discipline and Child Behaviour as Predictors of Children's Expression of Prosocial Story Themes Two Years Later* 168
Table 28 *Linear Model of Early Parental Discipline and Child Behaviour as Predictors of Children's Narrative Coherence Two Years Later* 171
Table 29 *An Overview of Widely Used Measures of Parental Sensitivity* 191
Table 30 *Intra-Class Correlation Coefficients for the Ainsworth and Erickson Sensitivity Scales* 206
Table 31 *Recruitment Pathway of Participants Disaggregated by Treatment Arm* 219
Table 32 *Recruitment Location of Participating Families Disaggregated by Treatment Arm* 219
Table 33 *Sociodemographic Characteristics of the Full HSHS Sample and the Subsample Included in the Current Analysis* 221
Table 34 *Descriptive Statistics of Study Variables for the Full HSHS Sample and Subsample Included in the Current Analyses* 225
Table 35 *Bivariate Correlations Between Measures of Parental Sensitivity, Child Behaviour, and Children's Narrative Representations* 226
Table 36 *Linear Model of Predictors of Narrative Coherence* 230
Table 37 *Linear Model of Predictors of Behavioural Dysregulation* 232
Table 38 *Linear Model of Predictors of Children's Expression of Disruptive Story Themes* 234
Table 39 *Linear Model of Predictors of Children's Expression of Prosocial Story Themes*. 236
Table 40 *Linear Model of Predictors of Five-Month (Post-Intervention) Parental Sensitivity* 239
Table 41 *Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Supportiveness (Erickson Subscale)* 240
Table 42 *Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Intrusiveness (Erickson Subscale)* 241

Table 43 *Baseline and Five-Month Parental Supportiveness and Intrusiveness by Trial Arm* 242
Table 44 *Linear Model of Predictors of Post-Intervention (Five-Month Follow-Up) Child Behaviour (CBCL)* 246
Table 45 *Linear Model of Predictors of Five-Month (Post-Intervention) Parental Sensitivity* 247
Table 46 *Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Supportiveness (Erickson Subscale)* 248
Table 47 *Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Intrusiveness (Erickson Subscale)* 249
Table 48 *Baseline and Five-Month Parental Supportiveness and Intrusiveness by Trial Arm* 251
Table 49 *Sociodemographic Information for Children's Secondary Caregivers* 335
Table 50 *Intraclass Correlation Coefficients for Story Stem Coding* 358
Table 51 *Sociodemographic Characteristics of Samples by Inclusion in Chapter* 359
Table 52 *Descriptive Statistics of Measures of Child Behaviour, Parental Psychopathology and Parental Discipline of Samples by Inclusion in Chapter* 361
Table 53 *Bivariate Correlations Between Teacher-Rated Child Behaviour and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment* 364
Table 54 *Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Expression of Disruptive Themes Two Years Later* 368
Table 55 *Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Observed Behavioural Dysregulation Two Years Later* 369
Table 56 *Linear Model of Predictors of Post-Intervention (Five-Month Follow-Up) Child Behaviour (CBCL) for Full HSHS Sample* 370

List of figures

Figure 1. Key concepts and questions addressed in this chapter and the ways in which these feed into the broader aims of this thesis. 20
Figure 2. An overview of my contributions to the HSHS trial and my embedded PhD research.. 51
Figure 3. Flow of all HSHS participants through trial. 55
Figure 4. Mean scores of coded behaviours using summed scores across the three story stems. 103
Figure 5. Mean scores of coded thematic content disaggregated by story stem. 104
Figure 6. Children's narrative coherence disaggregated by age and gender. 116
Figure 7. A visual depiction of the ways in which children's behaviour, parenting exposures, and internal representations may be related to one another. 140
Figure 8. Timing of data collection for the examination of longitudinal associations between early child behaviour and parenting environment and later narrative representations. 142
Figure 9. The CONSORT flow diagram for participant data included in the current analyses... 154
Figure 10. Visual depiction of some of the defining features of parental sensitivity outlined in the literature. 187
Figure 11. Timing of data collection to explore links between early parental sensitivity and children's later behavioural and representational outcomes. 203
Figure 12. Excerpt of coding notes taken from a coding case completed in the current study.... 208
Figure 13. Simple mediation model: The mediating effect of post-intervention parental sensitivity on children's narrative coherence..... 215
Figure 14. Simple mediation model: The mediating effect of post-intervention parental sensitivity on children's behaviour problems. 215
Figure 15. The CONSORT flow diagram for participant data included in the current analyses. 218
Figure 16. Simple mediation model of intervention effects on children's narrative coherence mediated by parental sensitivity at 5-month follow-up. 244
Figure 17. Mediation model of intervention effects on children's post-intervention behaviour mediated by parental sensitivity at 5-month follow-up. 252
Figure 18. Summary of the key components of the discussion chapter. 268
Figure 19. A depiction of the ways in which key findings of each study's chapter feeds into the overarching findings summarised in this section. 269
Figure 20. Distribution of mean disruptive theme scores split by child gender. 365
Figure 21. Distribution of mean prosociality scores split by child gender. 365
Figure 22. Distribution of mean narrative coherence scores split by child gender. 366
Figure 23. Distribution of behavioural dysregulation scores split by gender. 366
Figure 24. Means scores of prosociality by age of child. 367
Figure 25. Means scores of narrative coherence by age of child. 367

List of appendices

Appendix A *Ethical Approvals for the Healthy Start, Happy Start Study* 329

Appendix B *Sociodemographic information of children's secondary caregivers* 335

Appendix C *Phase 1: Participant Information Sheet and Consent Form* 336

Appendix D *Phase 2: Participant Information Sheet and Consent Form* 338

Appendix E *Administration manual for the doll's house play story stem battery* 344

Appendix F *Coding manual for the doll's house play story stem battery* 355

Appendix G *Intraclass correlation coefficients for story stem coding* 358

Appendix H *Participant sociodemographic information and variable descriptives disaggregated by chapter subsamples* 359

Appendix I *Data testing univariate assumptions of normality* 362

Appendix J *Bivariate correlations between teacher-reported child behaviour and parent-related variables* 364

Appendix K *Distributions of story stem constructs split by child gender* 365

Appendix L *Distributions of story stem constructs by child age* 367

Appendix M *Secondary analysis: Early child behaviour as a predictor of children's disruptive themes* 368

Appendix N *Secondary analysis: Early child behaviour as a predictor of children's observed dysregulation* 369

Appendix O *Post hoc analysis: Group allocation as a predictor of children's behaviour at the 5-month follow-up for full HSHS sample* 370

Acknowledgements

To Dr Christine O'Farrelly and Professor Paul Ramchandani. There is no way a few lines in an acknowledgement section will allow me to express how immensely grateful I am to have had you both as supervisors, mentors, and unwavering sources of support over the last five years. For your expertise and guidance but, more than that, your kindness, encouragement, generosity, and belief in me and this work. We did it!

To the Imperial College London President's PhD Scholarship scheme and the generosity of alumni donors that makes it possible. This scholarship opened up a life and opportunities far beyond what I could have imagined.

To everyone involved in the design, data collection, programme delivery, and analysis of the Healthy Start, Happy Start study. Particular thanks go to the Project Management Group, Trial Steering Committee, Data Monitoring and Ethics Committee, Imperial Clinical Trials Unit, and the Patient and Public Involvement Group. To the trial sponsors, Imperial College London, and funders, the NIHR HTA programme (13/04/33). Happy Start not only formed the foundation of this thesis but also my entire working and research life for five and a half years. I feel very lucky to have been a part of it.

To all those in the Division of Psychiatry, Imperial College London who contributed to and supported Happy Start and, in turn, this thesis. An extra special thanks to Ellen Grimås and Holly Mattock who were so central to the collection of the two-year follow-up data and the maintenance of the HSDT. Thank you to my internal examiners, Professors Elena Garralda Hualde and Christopher Millett, for your expertise and guidance in the foundational stages of this work. To Professor Marian Bakermans-Kranenburg for the insightful training in parental sensitivity. To Professors Lynne Murray and David Oppenheim for valued thoughts on the coding of the children's data. A special thanks to Catherine Hatcher, Wei Luo, and Yanwen Wu for your contributions to coding. To my friends and colleagues in the PEDAL research centre, University of Cambridge who provided much-valued encouragement, feedback, and virtual lockdown working days.

To the Happy Start children, families, and dogs. This thesis does little to convey the joy and privilege it was to spend three years carting a suitcase filled with toys around, watching little toddlers transform into kind, sticker-loving, funny, bright, and wonderfully opinionated three- and four-year-olds. In research, hundreds of hours of fun, play, and learning often get translated into statistical output and tables filled with numbers. But the time I got to spend in

your homes helped me learn a whole lot more about child development, the importance and value of listening to children, and the tough and powerful job of parenting than any numbers on a screen. I am so grateful for your time and openness, and all it taught me.

Finally, to my family and friends, who are unlikely to ever see this (unless I now make thesis reading a compulsory part of dinner parties), but who deserve a monumentally special acknowledgement for bringing support, love, and perspective every step of the way. I truly could not have done this without you (particularly the whole global pandemic part!). Thank you for sticking with me. I cannot wait to *see* you and *hug* you!

CHAPTER 1: Behaviour problems, the parent-child relationship, and children's voices: An overview

1.1 Chapter rationale

In this chapter, I introduce concepts that are of key relevance to this thesis. The chapter begins with a summary of a set of theoretical frameworks that apply to aspects of the studies making up this doctoral research. Then, early behaviour problems, the parent-child relationship, and the inclusion of children's voices in research practice are discussed in turn. The chapter ends with an outline of the overarching aims of the current thesis.

1.2 Introduction

The central aims of this thesis were to explore the integration of children's perspectives into investigations of the presentation, development, and intervention of early behavioural difficulties in very young children, and to investigate the influences of parental discipline, parental sensitivity, and a parenting intervention on children's outcomes.

The studies making up this work are characterised by three overarching themes: behaviour problems in young children, the role of parenting and the parent-child relationship in children's earliest years of development, and the integration of very young children's (under five years old) perspectives into our understanding of these trajectories. Chapter two introduces the Healthy Start, Happy Start study, the trial in which the current set of studies were embedded. Chapter three investigates the implementation of a brief play-based narrative measure of children's own perspectives with a sample of three- and four-year-olds. Chapters four and five focus more closely on the role of parenting on children's longer-term internal and external functioning. In Chapter four, associations between children's narrative representations and their early experiences of behaviour and parental discipline are explored, before effects of a video-feedback parenting intervention on children's narratives are examined. Chapter five extends upon this focus by examining early parental sensitivity as a predictor of children's internal functioning, as well as its role as a mediator in the intervention effects on children's outcomes.

This introductory chapter provides a broad overview of the theory, literature, and gaps in our understanding of the core concepts underlying the aims of this thesis. Subsequent chapters then consider relevant theory and concepts in finer detail, alongside the specific aims and rationale of each study.

Figure 1 provides an outline of the key themes and questions addressed in this chapter. In brief, I first consider the presentation, prevalence, and assessment of behaviour problems in young children, their associated difficulties, and the approach and timing of prevention and intervention efforts. Then I reflect upon the role and malleability of parenting as a key pathway through which positive child and family outcomes can be promoted. Finally, I discuss the value, challenges, and approaches to integrating the perspectives of very young children into studies of their behaviour and development. Taken together, these areas of research set the scene for the core questions and aims guiding this thesis.

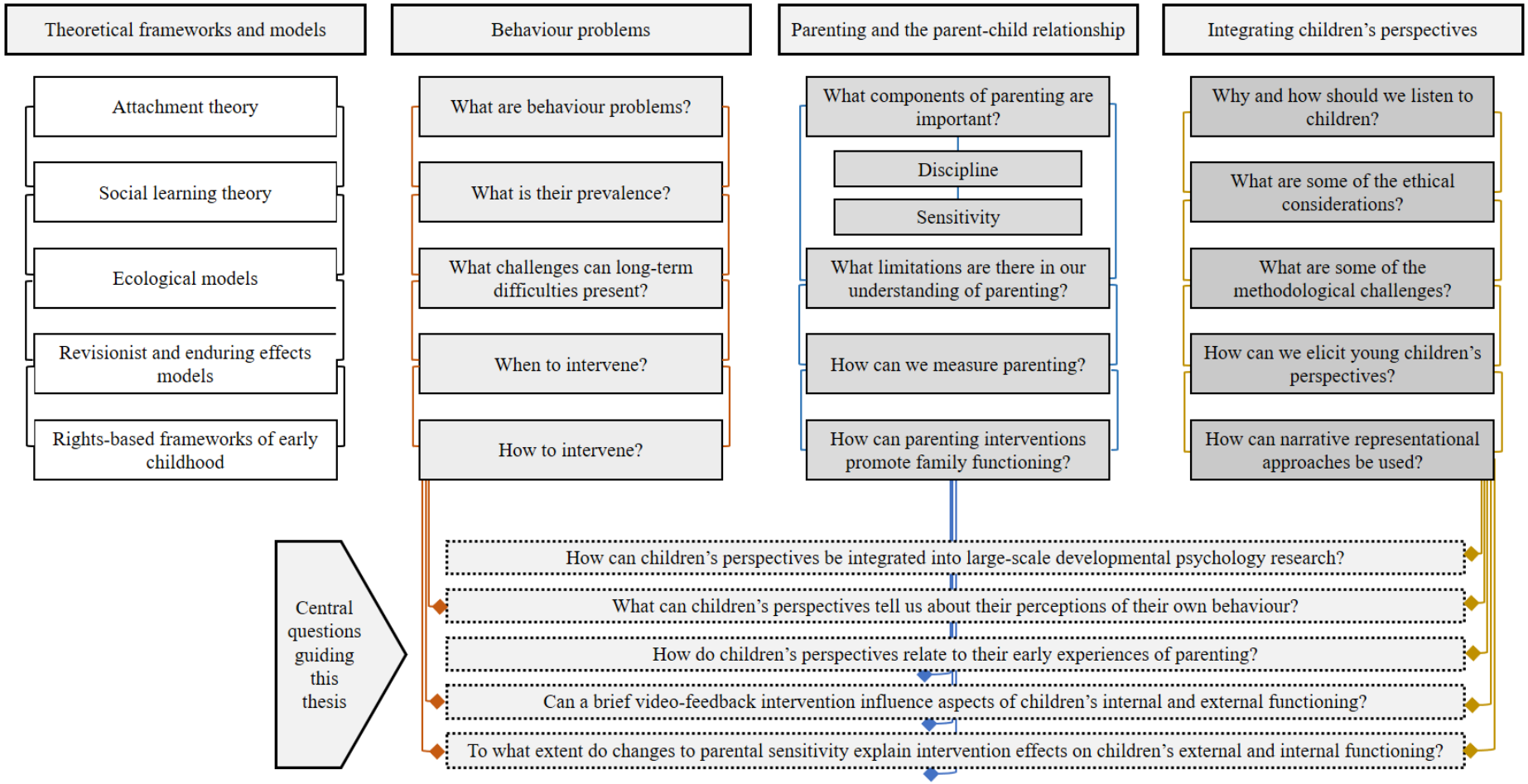


Figure 1. Key concepts and questions addressed in this chapter and the ways in which these feed into the broader aims of this thesis.

1.3 Theoretical frameworks and models

Before focusing on the specific components of this thesis, I first consider the strands of theory woven through the work as a whole. This research was not underpinned by a single theoretical framework. Instead, questions were generated and findings interpreted through different lenses of child development. A brief overview of several relevant theories is presented next.

1.3.1 Attachment theory

Applying ethological principles, early attachment theorists proposed that the propensity to form strong emotional bonds with caregivers is driven by a need for safety, security, and comfort (Bowlby, 1969/1982, 1973; Bretherton, 1987). Individual differences in attachment organisation are said to be detectable from the end of the child's first year of life (Grossmann & Grossmann, 2006). A caregiver who can accurately, adequately, and promptly respond to their child's cues (Ainsworth, Blehar, Waters, & Wall, 1978) is suggested to offer the child enough security to "explore their environment, take on challenges, engage in independent activities, and make discoveries" (Feeney & Woodhouse, 2016, p.827). Caregiver-child relationships characterised by secure mental representations of the attachment figure have been positively associated with children's social (Groh et al., 2014) and behavioural (Fearon, Bakermans-Kranenburg, van IJzendoorn, Lapsley, & Roisman, 2010; Groh, Roisman, van IJzendoorn, Bakermans-Kranenburg, & Fearon, 2012) development. Several parenting programmes aim to promote key components of secure attachment relationships in an effort to improve child and family wellbeing. However, insecure attachment is not synonymous with less optimal outcomes, nor is a secure attachment guaranteed to protect against behavioural challenges (O'Connor & Scott, 2007). It should also be noted that attachment security has been found to be only moderately stable across early childhood ($\kappa = 0.23$; Opie et al., 2020). Discussions of the caregiver-child relationship as conceptualised in attachment theory are seen throughout this thesis.

1.3.2 Social learning and coercion theories

Social learning theory (e.g., Bandura, 1977; Bandura & Walters, 1963) posits that child development is both directly and indirectly shaped by their real-life experiences. It is proposed that children's behaviour is shaped by observing and imitating the models around them, as well as the ways their own actions are responded to (O'Connor & Scott, 2007). Models of parenting that are rooted in social learning theory highlight the significance of

consistent discipline and the avoidance of 'coercive cycles' (Patterson, 1982). The current thesis considers the effects of a parenting programme that draws on coercion theory to promote a sensitive discipline style in caregivers.

1.3.3 Ecological models of development

Bronfenbrenner's theory proposes that child development occurs in the context of complex, reciprocal interactions between the child and multiple levels of their surrounding environment (e.g., Bronfenbrenner, 1979; Bronfenbrenner & Morris, 1998). In early versions of this theory, these complex environmental layers were described as the 'microsystem' (e.g., the child's direct interactions with caregivers or teachers), 'exosystem' (e.g., interactions between parents and teachers), and 'macrosystem' (e.g., attitudes, ideologies, and culture of social networks, neighbourhoods, and communities) (Bronfenbrenner, 1977). This thesis subscribes to the importance of person-context interrelatedness by focusing closely on the influences of children's immediate environs or 'microsystems' (i.e., exposures to parenting practices) on their behaviour and development. However, the work does not explicitly extend out into the 'exosystem' or 'macrosystem'.

1.3.4 Revisionist vs. enduring effects models of development

Whether children's earliest experiences leave a lasting or transient mark on their development has been conceptualised into two distinct models: the revisionist model and the enduring effects model (Fraley, Roisman, & Haltigan, 2013; Haltigan, Roisman, & Fraley, 2013; Raby, Roisman, Fraley, & Simpson, 2015; Roisman & Fraley, 2012). The revisionist model emphasises the malleability of development. It proposes that children's early experiences have a direct influence on their early development but that this effect dissipates over time due to the rich interplay of personal and social influences encountered through the life course (Clarke & Clarke, 2000). Under this model, associations between early caregiving and children's developmental outcomes are predicted to get smaller over time. In contrast, the enduring effects model suggests that children's early experiences continue to shape their development through the life course, via mechanisms established in the earliest years of life (Grossmann, Grossmann, & Waters, 2006; Sroufe, Coffino, & Carlson, 2010). This model anticipates that associations between early experiences and development will be relatively stable over time and – even where effects are small – consequential (Fraley et al., 2013; Raby et al., 2015). The studies in this thesis consider child development across two time points over two years, therefore, we are limited in our understanding of how these effects pattern over

time or endure in the long-term. Nonetheless, a central component of this research is the suggestion that early intervention could have lasting consequences for children's development, a proposal that ties into the enduring effects model. Conversely, it is also acknowledged that families are likely to benefit from staged supports through the family life course, with a recognition that effects of environmental enrichment may decline over time, in line with a revisionist framework.

1.3.5 Rights-based and democratic frameworks of early childhood.

The views of the child and childhood have shifted considerably in the last few decades. Acknowledgement and advocacy for children's rights now infiltrates political discourse and initiatives, policy frameworks, and – albeit to a lesser extent – research studies and publications (Hayes & Bradley, 2009). Under a rights-based framework, children have been positioned as visible, capable, and competent individuals who can offer valuable and original points of view (Lansdown, 2001). The formal endorsement of these rights through socio-political frameworks, including the United Nations Convention on the Rights of the Child (UNCRC; 1989), ratified by all UN member states except the USA, and the UN Commission (2005), pushed the need to value children's perspectives in decision-making to the global fore. Article 12 of the UNCRC states that "parties shall assure to the child who is capable of forming his or her own views the right to express those views freely in all matters affecting the child…[and have] the opportunity to be heard" (United Nations Convention on the Rights of the Child, 1989, pg. 6). Especially significant for this thesis is the fact that no lower age limit is imposed on the need to support children in exercising their right to participate in the expression of their views. More recently, the adoption of the new EU Strategy on the Rights of the Child (European Commission, 2021) marked a commitment to "promote and improve the inclusive and systemic participation of children" to support them in shaping and evaluating decisions that affect their lives. The WHO-UNICEF-Lancet Commission (Clark et al., 2020) also emphasised the importance of children's involvement by stating that strategies to "improve child health and wellbeing cannot succeed unless they are truly multisectoral" (p. 608). Despite these strong endorsements, there remains a translational gap between the legal and political commitment to the principle and the more widespread practice of facilitating children's meaningful and effective participation (Lundy, 2007). The studies in this thesis are underpinned by the belief that children's involvement in research can bring significant value to our understanding of their development.

In sum, throughout this thesis these diverse frameworks helped inform my conceptualisations of child development and parenting, the interpretation of findings, and discussions around possible implications. Next, I focus closely on the three core components of this research in turn: children's behaviour problems, parenting, and children's voices.

1.4 Behaviour problems

In this section, I consider our understanding of the presentation and prevalence of behaviour problems, reflect upon their associated challenges, and discuss ways in which children and families can be supported with these difficulties.

1.4.1 What are behaviour problems?

All children misbehave from time-to-time. Rule breaking, disobedience and moments of 'pushing the boundaries' are typical components of healthy child development. For most children these behaviours pass in time or are only shown occasionally. Yet, for some children, these problems can become so frequent, severe, and persistent that they represent a pattern of behaviours that are of clinical significance (Gutman, Joshi, Khan, & Schoon, 2018).

Behavioural difficulties are often separated into internalising and externalising problems. Internalising problems are characterised by children channelling their emotional responses inwards, which can present as withdrawn or attention-seeking behaviours, an over-dependence or clinginess to others, and low self-esteem (McCulloch, Wiggins, Joshi, & Sachdev, 2000). Conversely, externalising problems are characterised by under-controlled, outward-facing emotions. The nature of externalising behaviours can vary according to the child's age (Gutman et al., 2018). In young children, we might see tantrums, disobedience, destructiveness and physical aggression, while in older children these behaviours can present as bullying, stealing, and lying. In adolescence, these problems can escalate into violence, truanting, substance misuse, and antisocial behaviours like shoplifting or vandalism (Green, McGinnity, Meltzer, Ford, & Goodman, 2005; Gutman et al., 2018). Externalising behaviours that are related to attentional difficulties, rather than conduct, are characterised by "pervasive, developmentally excessive, and impairing levels of overactivity, inattention, and impulsivity" (Sayal, Prasad, Daley, Ford, & Coghill, 2018, p.175).

Comorbidities of emotional and behavioural problems are commonly seen (Angold & Costello, 1993; Lahey, Loeber, Burke, Rathouz, & McBurnett, 2002; Lilienfeld, 2003). A recent meta-analysis from Vasileva, Graf, Reinelt, Petermann, and Petermann (2021; 10 studies; N = 18,282 children) found that one third of children aged one to seven years with

poor mental health fulfilled the criteria for more than one diagnosis. Wichstrøm et al. (2012) reported that, in a sample of 1,250 four-year-olds, comorbidities between emotional and behavioural disorders ranged from 21 to 85%. Issues related to each set of problems may play a causal role in the origins of the other. For instance, chronically low mood may impair an individual's capacity to be concerned about adverse consequences of their actions (Capaldi, 1991) and/or externalising behaviours may give rise to a recurrent state of anxiety in individuals (Keiley, Lofthouse, Bates, Dodge, & Pettit, 2003). However, conversely, it has been suggested that symptoms related to anxiety could be protective for children experiencing conduct difficulties (Walker et al., 1991). The complexities of these possible co-morbidities are not addressed in this thesis but are important to recognise when more widely considering the landscape of mental health in infants and pre-schoolers.

1.4.2 What is the prevalence of these difficulties?

Problems associated with attention and conduct are some of the most common symptoms of mental health diagnoses in children (Maughan, Collishaw, Meltzer, & Goodman, 2008; Sadler et al., 2018). In children aged five to sixteen years, an estimated 6% meet the diagnostic criteria for conduct disorders (Green et al., 2005) and 2-7% for hyperkinetic disorders (Sayal et al., 2018). Similar numbers of children exhibit behaviour problems that, while not severe enough to meet the thresholds for clinical diagnosis, are still disruptive and distressing for the individual and those around them (Gutman et al., 2018; Sayal et al., 2018).

Prevalence of these problems increases for children facing additional adversities. Large survey data of children aged five to ten years old in England indicated that 36% of looked-after children displayed clinically significant conduct disorders compared to 5% in private households, while 11% had hyperkinetic disorders, compared to 2% in private households (Meltzer, Gatward, Corbin, Goodman, & Ford, 2003a). Experiences of abuse in childhood may also be linked to these difficulties. A systematic review from Maniglio (2014) reported that an estimated 27% of participants with conduct disorder had experienced childhood sexual abuse, compared to 11% of controls. Socioeconomic factors are also of importance, with Green et al. (2005) reporting that 52 to 58% of children with conduct and hyperkinetic disorders lived in households with gross weekly household incomes under £300, compared with 23 to 34% of typically developing children.

Investigations of the prevalence of mental health difficulties in very young children (~5 years and under) are few (Lyons-Ruth et al., 2017) and thus our understanding of the epidemiology of behaviour problems in these children remains further behind that of older children (Egger & Angold, 2004). A key challenge in accurately identifying these problems in toddler- and preschool-aged children is the rate of development that occurs during this life stage. Children are rapidly developing, acquiring, and testing out key skills, including the capacity to sustain attention, inhibit behaviour, interact with others, and process and communicate emotions. Defining clear clinical thresholds for these behaviours, while accounting for typical developmental variation, is difficult. Similarly, care must be taken to distinguish between transient individual differences in children's development and clinically significant difficulties (Zeanah Jr & Zeanah, 2019). For instance, while tantrums are an expected feature of toddlerhood, violent, prolonged, and/or frequent (e.g., daily) tantrums are not (Biedzio & Wakschlag, 2019; Wakschlag et al., 2012). Further to this, in the early years especially, caregivers and their interactions with their children play a key role in supporting children's regulation of emotions and behaviour (Lincoln, Russell, Donohue, & Racine, 2017). Because of this, it can sometimes be unclear whether children's disruptive behaviours are a manifestation of their own individual difficulties or an expression of challenges in the caregiver-child relationship (von Klitzing, Döhnert, Kroll, & Grube, 2015).

Using the assessment tools that are available for children in this younger age range, the estimated prevalence of diagnosable mental health difficulties in preschool-aged children (two- to four-year-olds) is as follows: any diagnosable mental health disorder: 5.5%; emotional disorders: 1.0%; behavioural disorders: 2.5%, hyperactivity disorders: 0.5% (Mandalia et al., 2018). It is worth noting though that this survey used operationalised diagnostic criteria in their assessments and thus these figures only represent those children with particularly significant levels of difficulties. The authors advise exercising caution when interpreting these findings, proposing that an increase in prevalence may have been seen had the data been obtained through multiple informants (e.g., teachers), as was the case for the surveyed five- to ten-year-olds where rates were markedly higher (5.5% vs. 9.5%). In a review of mental health difficulties in pre-schoolers (two- to five-years-olds), Egger and Angold (2006) reported prevalence rates of behaviour disorders at 9 to 15% and emotional disorders at 9 to 11%, while Skovgard (2010) reported prevalence of 2.4% for attentional difficulties and 4.3% for behavioural and emotional disorders (considered together) in children aged only one-and-a-half years old.

1.4.3 What challenges can long-term difficulties present?

Disruptive behaviour disorders can present multiple challenges across the lifespan, with Gutman et al. (2018) emphasising that "no other common childhood condition is associated with such far-reaching and pervasive consequences" (p.3). A consideration of these consequences makes a clear case for the importance of early and accurate diagnoses and effective interventions. Throughout the writing of this thesis, I have aimed to strike a balance between communicating the important need to better understand child development and the role parents can play in their children's lives, without writing in too deterministic a way with regards to children's long-term outcomes. The forthcoming section communicates several important aspects of development that have been found to be associated with enduring behavioural difficulties. However, they should be considered with an appreciation of the individuality in children's experiences, and the potential for more positive outcomes.

Investigations into the developmental trajectories of adult psychiatric difficulties identify early-onset conduct disorders as an influential factor. Using New Zealand birth cohort data (N = 1,037), Kim-Cohen et al. (2003) reported that 25 to 45% of adults with a psychiatric disorder had a diagnosable conduct disorder in childhood (measured at 11 to 18 years). Importantly, juvenile conduct disorder preceded all adult disorders measured, including anxiety, depression, substance dependence, and eating disorders. Thus, identifying and targeting behaviour problems in childhood represents a vital prevention opportunity for adult mental health adversity.

Education-wise, using UK survey data from 7,977 children, Green et al. (2005) found that children with hyperkinetic and conduct disorders were especially likely (12% and 15% respectively) to have been excluded from the classroom three or more times by the age of 16, compared to 5% with emotional disorders, and 0-1% with no disorder. Importantly, Meltzer, Gatward, Corbin, Goodman, and Ford (2003b) reported that 26-37% of children with persistent conduct/hyperkinetic problems over three years had been excluded, compared to 10-12% of children who exhibited these difficulties in the first survey but whose problems had declined by the three-year follow-up survey (N = 2,938). This demonstrates the disparities that can emerge when problems endure. Effects on higher education are mixed, but Kuriyan et al. (2013) reported that, compared to 48% in the control group, only 15% of young adults with ADHD held a degree, and were 11 times more likely to be unemployed.

As well as the effects on individual families, there are high economic and societal costs associated with behaviour problems. Using a whole-population cohort of 120,000 children, Spencer et al. (2005) reported that children with moderate to severe conduct disorder were seven times more likely to be on the child protection register. In a 25-year longitudinal study, Fergusson, Horwood, and Ridder (2005) found that 37% of participants with high conduct problems aged seven to nine became parents as teenagers, compared to 12% in the low problems group.

The Sainsbury Centre for Mental Health (2009) estimated that about 80% of those with experience of the criminal justice system had conduct problems in childhood. Gutman et al. (2018) predicted that the individual cost of childhood conduct disorder is £5,500 a year, with 60% of this falling on the education sector. Scott, Knapp, Henderson, and Maughan (2001) projected individual public sector costs up to the age of 28 to be £100,267 for those with conduct disorder, £34,832 for those with threshold difficulties, and £10,630 for those with no such problems. Individual lifetimes costs for people with early conduct disorder have been estimated at £280,000, with sub-threshold difficulties projected at £90,000 (Friedli & Parsonage, 2007).

In sum, embedded and enduring behavioural difficulties have potentially far-reaching effects for children, families, and wider society. When considered together, there is a strong case for the implementation of effective prevention and intervention strategies.

1.4.4 When to intervene?

Families with children aged three to five years are increasingly being offered access to treatment for established behaviour problems (Petitclerc & Tremblay, 2009). This raises key questions around whether waiting until the preschool and early education years, when children's behavioural difficulties might already be embedded, is the most optimal age for intervention. Intervening in the zero-to-three period may be especially powerful given the rapid rate of synapse production, heightened neural plasticity, and children's increased sensitivity to enriched environments in these years (Hannon, 2003; Winston & Chicot, 2016). If families most likely to benefit are effectively identified, the costs associated with early intervention programmes can be largely offset by the high rate of return that comes with reducing the impact of these problems early (Doyle, Harmon, Heckman, & Tremblay, 2009). This economic return is especially vital for over-stretched public health services like the NHS.

Researchers have emphasised the extra care and caution that must be applied when identifying 'disordered behaviours' in very young children. I have already discussed the inherent challenges of reliably assessing children during this time of rapid developmental change, and there are concerns around stigmatising young children if given a diagnosis (Carter, Briggs-Gowan, & Davis, 2004; Scott & Gardner, 2015). Equally though, failure to identify early behaviour problems and provide appropriate support may not only exacerbate difficulties but potentially diminish parents' sense of efficacy and confidence in parenting if their concerns are dismissed (Carter et al., 2004). Given that intervention in later childhood is frequently more costly and often cannot achieve the results that early intervention can (Allen, 2011), ensuring families have access to high-quality and evidence-based support in the earliest years of their children's lives is critical. This need has been recognised in global policies that champion the immense potential of early intervention (Allen, 2011; Department of Health and Social Care and Department for Education, 2017) and we have seen an increase in the development of programmes specifically designed for use with toddler-aged children (Bakermans-Kranenburg, van IJzendoorn, & Juffer, 2003).

1.4.5 How to intervene?

Identifying and understanding the risk and protective factors that underlie behavioural difficulties can allow for the development of reliable and effective intervention and prevention programmes. McLennan and Offord (2002) outlined that targets for successful intervention must have a high attributable risk to the outcome, be plausibly causational in their influence (based on empirical data or robust theoretical thinking), amenable to change, and identifiable through simple and accurate screening. Interventions must also be feasible to be delivered at scale and acceptable to key stakeholders.

We know that no single intervention can act as a 'magic bullet' for children's optimal development or a panacea for the varied challenges facing families (Allen, 2011; Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2017a). Instead, families are likely to benefit from a range of well-implemented, continually iterated upon, evidence-based programmes, selected based on the family need at that time (Pingault, Rijsdijk, Zheng, Plomin, & Viding, 2015). One significant factor – that forms a central mechanism of change for early behaviour problems, and is an acceptable and effective target for intervention – is parenting (Gardner & Shaw, 2008). Learning and practicing important parenting skills that adequately meet children's needs can uniquely position parents as one of the best-placed supports for their children's difficulties (Gardner & Shaw, 2008). It has been stated that it is what parents do, rather than who they are

or where they come from, that is especially important for children's development (Allen, 2011; Gardner, Hutchings, Bywater, & Whitaker, 2010).

Given this, in the next section I consider the ways in which parenting can contribute to children's development, aspects of parenting that may be especially noteworthy in the developmental trajectories of behaviour difficulties, approaches to measuring parenting, and ways positive parenting practices can be promoted to enhance children's outcomes.

To conclude this section, undertaking work to better understand the predictive and protective factors of behaviour problems in very young children is essential. It holds the potential to enhance the approaches undertaken to support families at the first signs of difficulties, making long-term differences to children's lives. Further to this, utilising novel methods and tools in such work may help to address existing measurement issues in assessing very young children during a dynamic stage of their development.

1.5 Parenting and the parent-child relationship

Understanding the ways in which parenting and the parent-child relationship shapes child development has been a significant source of inquiry for researchers, clinicians, policymakers, and families themselves for decades (O'Connor & Scott, 2007). A burgeoning evidence base suggests that the quality of parent-child interactions is a key influence on children's outcomes and a promising target for intervention (Britto et al., 2017; Donelan-McCall, 2017). In this section, I consider two important components of parenting and examine key approaches to measuring and promoting parenting practices. Note, the term 'parent' is used to describe any caregivers with primary responsibility of a child.

1.5.1 What components of parenting are important for children's behavioural development?

The field of parenting research encapsulates a diverse and complex array of characteristics related to family life. Here, I specifically focus on two components of parenting that are of central relevance to this thesis: parental discipline style and parental sensitivity.

1.5.1.1 Parental discipline

Caregivers must find ways to support their children as they learn to regulate their behaviour and adhere to familial and societal expectations and values. Discipline can be a defining component of these processes. Appropriate and effective consequences can be used

to elicit compliance, help children learn rules, and prevent or correct misbehaviour (Socolar & Stein, 1996). The approach of caregivers engaging in reactive (i.e., in response to misbehaviour) or proactive (i.e., for encouragement of desired behaviour) discipline has been found to be both more and less conducive to children's development (Gershoff & Grogan-Kaylor, 2016; Larzelere, 2000). 'Effective' discipline has been associated with lower behaviour problems (Choe, Olson, & Sameroff, 2013), a greater capacity for self-regulation (Karreman, van Tuijl, van Aken, & Deković, 2006), and higher levels of prosociality (Knafo & Plomin, 2006), while 'ineffective' or 'dysfunctional' discipline has been linked to lower emotion regulation and higher internalising and externalising difficulties (Chang, Schwartz, Dodge, & McBride-Chang, 2003; O'Leary & Vidair, 2005; Parent, McKee, & Forehand, 2016). An alternative delineation is offered by Diana Baumrind's work exploring power assertion across permissive, authoritarian and authoritative parenting styles (Baumrind, 1966, 2012). Permissive parents may avoid asserting power or give in without following through on consequences, while authoritarian parents are likely to assert coercive power focused on the retention of hierarchy and characterised by punishment and threats. Authoritative parents take a firm, reasoned, negotiable approach, employing inductive strategies such as limit setting, reasoning, and explaining (Hoffman, 2001). Authoritative parenting styles are regarded as particularly optimal for children's social and behavioural development (Baumrind, 1991).

Components of children's disruptive behaviours, including tantrums and physical aggression, peak in the toddler and preschool years (NICHD Early Child Care Research Network, 2004; Potegal & Davidson, 2003; Tremblay et al., 2004; Wakschlag, Tolan, & Leventhal, 2010). Running parallel to this is a surge in caregivers' attempts to engage in strategies to reduce and correct these challenging behaviours. Better understanding the ways in which discipline style is associated with children's functioning, during such a salient phase of behavioural development, could afford us helpful insight into the forms of support families could most benefit from.

1.5.1.2 Parental sensitivity

Parental sensitivity has received a significant amount of interest in the parenting and child development fields. Using Ainsworth, Bell and Stayton's (1974) definition, parental sensitivity describes a caregiver's ability to notice, interpret and respond in an appropriate and prompt manner to their child's cues and signals. Under attachment theory, a caregiver's capacity for sensitive caregiving feeds into the child's internal perception of that caregiver (Juffer, Bakermans-Kranenburg, et al., 2017a). Experiences with multiple attachment figures

are believed to be internalised and generalised into a set of continually updated assumptions and expectations about how attachment (and wider social) relationships work (Society for Emotion and Attachment Studies, 2021). A child whose caregiver struggles to respond in a prompt and responsive way is considered to be more likely to form an attachment style that is marked by inconsistent or unresponsive caregiving behaviours (Donelan-McCall, 2017).

Sensitive and responsive parental behaviours are considered to be important (but not the only) mechanisms for children's social, emotional, and behavioural functioning (DePasquale & Gunnar, 2020; Teti, Cole, Cabrera, Goodman, & McLoyd, 2017). Meta-analyses revealed associations between sensitive caregiving and a secure attachment style (Atkinson et al., 2000; De Wolff & van IJzendoorn, 1997; Goldsmith & Alansky, 1987), while empirical research has established links to children's behaviour problems (e.g., Fearon et al., 2010; Miner & Clarke-Stewart, 2008), self-regulation (e.g., Eisenberg et al., 2001), social competence (e.g., Kochanska, 1997), and linguistic development (e.g., Tamis-LeMonda, Bornstein, & Baumwell, 2001). These findings situate parental sensitivity as a significant factor in family functioning that may be an important and enduring contributor to children's long-term development (Raby et al., 2015).

Before reflecting on key approaches to measuring parenting, I first acknowledge some broad limitations in our understanding of the parent-child relationship and reflect upon key aspects of family experiences that are overlooked or under-represented in the included literature.

1.5.2 What are some of the limitations in our understanding of the parent-child relationship?

The consideration of the existing literature and the findings of this thesis must be interpreted with an awareness of enduring gaps in our understanding within this field.

First, when considering the literature on parenting, it is necessary to keep in mind the relatively modest magnitude of effects found in the majority of studies. Collectively, this indicates that other important influences are exerting effects too. Further to this, from an ecological perspective, the parent-child relationship is embedded in a web of social factors that mean the impact of parenting on children's outcomes is likely to be dependent on the wider social context family experiences take place in. Finally, the largely correlational nature of the existing body of literature means there is less certainty about the direction of effects. Parenting should not be perceived as a unidirectional, top-down process but rather a

relationship influenced by children's individual characteristics (e.g., temperament, gender, age) too (O'Connor & Scott, 2007). Of particular relevance to this thesis, differences in children's behaviour have been found to elicit harsher and less warm parenting styles (Anderson, Lytton, & Romney, 1986; Barkley, 1988). Put simply, high parental sensitivity and low dysfunctional discipline alone are not exclusive conditions to children's happy and healthy development.

In this thesis I broadly discuss 'parenting' and the 'parent-child relationship'. However, important components of family experiences, including the unique influences of fathers, the couple relationship, and family systems (e.g., Foley, Álvarez, McCarthy, & Hughes, 2020; Rodrigues et al., 2021) are not captured in my conceptualisations and measurements of parenting. Historically, 'mothering' was largely assumed to be synonymous with 'parenting' (Pleck, 2012). Despite an increased recognition of the importance of paternal involvement, there remains a persistent under-representation of fathers in parenting research, parenting programmes, and the provision of services to children (Barker, Iles, & Ramchandani, 2017; Cabrera, Volling, & Barr, 2018). Additionally, other sources of caregiving are unaccounted for including custodial or multigenerational care provision from grandparents or other kinship carers (Dunifon, Near, & Ziol-Guest, 2018; Sadruddin et al., 2019) and childcare provided outside of the home (McCarthy & Morote, 2009; Sims, Guilfoyle, & Parry, 2006). Also, sociocultural inequalities related to poverty, unemployment, ethnicity, and social mobility, among others, can have highly adverse impacts on families' capacity to parent in the way they would like to (Morrison, Pikhart, Ruiz, & Goldblatt, 2014). The presence and complexities of these inequalities remain mostly unaddressed in this thesis but are important to identify when considering the impacts of parenting on children's lives.

Finally, the parenting field is limited by the fact its understanding of parent-child interactions is overwhelmingly built on findings drawn from studies of highly-educated, nuclear families with secure incomes living in settings in North America, Western Europe, and Australia (Henrich, Heine, & Norenzayan, 2010; Keller et al., 2018). Claims that Ainsworth's original definition of sensitive caregiving is universally valid (see Mesman et al., 2018) have recently been disputed. Keller et al. (2018) argue that this definition does little to account for the cultural specificity of caregiving that hinges on variance in socialisation goals and developmental needs across and within cultures. More work needs to be done to ensure that caregiver and child behaviours regarded as key for children's development are culturally relevant and sensitive (Mesman, 2021; Mesman et al., 2018). Thus, in the forthcoming set of

studies – drawn from a sample of families living in the UK – limitations to the relevance and applicability of findings to wider contexts must be kept in mind.

1.5.3 How can we measure parenting?

Variation in approaches to assessing parenting capture different characteristics, dimensions, or accounts of parental behaviour or style. Here I focus on measuring parenting through questionnaires and observation.

1.5.3.1 Questionnaires

Questionnaire measures offer a helpful way to capture data quickly and efficiently, particularly in large samples, increasing statistical power and generalisability (Zaslow et al., 2006). Questionnaires can be particularly useful as part of a wider assessment battery, allowing for insight into multiple parent-related behaviours, with a lower level of burden for families and researchers (Ramey, 2001). Self-reported data has allowed the field to gain an in-depth understanding of parental behaviours and their associations with children's development across diverse settings and contexts. However, asking parents to report on their own behaviour can introduce measurement error in terms of disparities in individual participants' understanding and standards (Clayman & Wissow, 2004). Also, both conscious and unconscious bias related to social desirability and a perceived intrusiveness or risk in disclosure of truthful answers to sensitive questions (e.g., use of physically violent discipline practices) influence reporting (Morsbach & Prinz, 2006; Tourangeau, Rips, & Rasinski, 2000).

1.5.3.2 Observation

Standardised observational approaches are often pitched as a 'gold standard' form of assessment (Hawes & Dadds, 2006), with the potential to offer a more reliable, objective, and complex insight into parents' behaviour patterns (Zaslow et al., 2006). These techniques can be particularly beneficial for measuring possible behavioural changes brought about by interventions, as caregivers may overestimate changes in their own behaviour (Patterson, 1982). Observational approaches may be especially valuable in the measurement of parental sensitivity as caregivers can only report on signals they notice, not those they miss or misinterpret; "one can hardly be aware of being unaware" (Mesman, 2021, p. 116). Videotaped observations in particular can bring several advantages including the ability to establish inter-rater reliability, repeatedly review single moments, undertake multi-layered assessments of behaviour in the same interaction, and engage in discussions between

observers (Mesman, 2021). However, these in-depth and rigorous approaches can be expensive, time-consuming, and reliant on the acquisition of a refined set of skills, rendering such techniques largely inaccessible to low-resourced contexts. Finally, in this field observations are typically undertaken on interactions lasting minutes. While scales applied to such short observations have exhibited promising psychometric properties (e.g., Aspland & Gardner, 2003; Mesman & Emmen, 2013), these relatively brief moments are unlikely to fully capture the richness and variety in caregiver-child moments day-to-day. The presence of a researcher may also elicit a level of 'observer reactivity' from participants (Gardner, 2000).

1.5.3.3 Measurement of parenting in this thesis

In this research, parental discipline was measured using a self-reported questionnaire and parental sensitivity was measured using observations of parent-child interactions. Obtaining data on two components of parenting has the potential to provide a richer understanding of children's experiences of parenting. However, each approach brings unique advantages and disadvantages that will be detailed in subsequent chapters. Having been fortunate enough to collect this data first-hand, I can appreciate the fine balance that must be struck between accessing 'gold standard' data and ensuring families are not fatigued or burdened by their participation. The measures included in this thesis were components of wider research assessments conducted with caregivers and their one- to five-year-olds that lasted approximately two and a half hours. Using a mix of gold standard (i.e., clinical interviews, observations) and quick and simple (i.e., questionnaires) measures ensured families were not over-burdened, something that may well have contributed to the high degree of retention seen in the forthcoming set of studies (85-95%).

Next, I reflect upon ways key components of parenting can be promoted with the intention of improving child and family functioning.

1.5.4 How can parenting interventions promote family functioning?

Evoking effective and reliable change in the quality of caregiving children receive, in ways that accommodate for individual, familial, and contextual factors, is a significant challenge facing the field. Nonetheless, there is a larger empirical evidence base for parenting interventions than any other psychological intervention for children's mental health (Scott & Gardner, 2015). These programmes are largely underpinned by social learning and attachment theories and often focus on promoting positive parenting practices and/or reducing negative ones (Duncombe, Havighurst, Holland, & Frankling, 2012). Table 1 outlines a selection of

parenting programmes that have been trialled with families, with a focus on their implementation and effectiveness in UK settings.

There is convincing evidence that parenting interventions offer a powerful tool to enhance aspects of family functioning. A meta-analysis (N = 54 studies) of parenting interventions in families with children aged three to ten years revealed moderate reductions in children's behaviour (0.40-0.69 SD) (National Institute for Health and Care Excellence, 2013). Gardner et al. (2019) also reported positive effects on children's disruptive behaviours (d = -.47) in a meta-analysis of 154 trials testing 50 parenting interventions (N = 13,387 children; mean age: 5.3 years). Two umbrella meta-analyses of 11 and 26 meta-analyses respectively found positive effects of parenting interventions on parenting characteristics (SMD = 0.52-0.60) and children's externalising difficulties (SMD = 0.45) in children aged 0 to 18 years (Mingebach, Kamp-Becker, Christiansen, & Weber, 2018; Weber, Kamp-Becker, Christiansen, & Mingebach, 2019). Improving parenting quality may not only have positive impacts on children's behaviour but also bring significant savings, both to individual families and wider society. A recent longitudinal study first assessed children when they were four to six years old and again in early adolescence, and found that sensitive early parental care predicted significantly lower costs, in terms of family expenditure and education, health, social, and justice service costs, through to early adolescence (most sensitive quartile mean £1,619, least sensitive quartile mean £21,763; Bachmann, Beecham, O'Connor, Briskman, & Scott, 2021).

We have seen considerable advances in the development and evaluation of parenting interventions in recent years. However, interventions that appear to work when delivered by research therapists, in tightly-controlled, well-resourced, clinical research studies, widely fail to replicate when implemented in 'real-world' public health service settings delivered by non-specialised clinicians (Scott & Gardner, 2015). There also remains a lack of clarity around which programmes work best, for whom, and under what conditions. Although progress has been made, there remains a scarcity of interventions designed for delivery in families with toddler- and preschool-aged children, despite the evidence discussed earlier that effective early interventions may offer the greatest return, both for families and society.

One intervention that has been widely tested with families of very young children is the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2008b, 2017b; Mesman et al., 2008). This programme forms the focus of the next section.

Table 1

Overview of Parenting Interventions Tested with Families of Young Children

Intervention	Child age	Key components of delivery	Evidence base (with a focus on intervention delivery in the UK)
Family Nurse Partnership (adapted from the Nurse Family Partnership; Olds, 2006)	Pregnancy-2 years	Involves up to 64 home visits. Nurses discuss health, maternal role, and access to services, to target prenatal health-related behaviours, sensitive caregiving, and early parental life course.	In a pragmatic trial (Robling et al., 2016; N = 1,645) of expectant mothers aged ≤19 years, following a mean of 39 home visits from family nurses from pregnancy until two years of age, no evidence of effects on smoking during pregnancy, infant birthweight, emergency hospital admissions, or timing of second pregnancies were found. At follow-up (2-6 years), there was evidence of greater school readiness in children in the trial arm (Robling et al., 2021).
Incredible Years® (IY; Webster-Stratton, 2001; Webster-Stratton, 2008)	0-13 years	Group-based, uses video clips of model families. Families practice techniques through role play and homework.	A meta-analysis found positive effects for children's disruptive (d = .27) and prosocial (d = .23) behaviours (Menting, de Castro, & Matthys, 2013; N = 50; mean age: 3.0-9.2 years). However, a pragmatic trial of the toddler version with families of 1-3 year olds (N = 89; Hutchings, Griffith, Bywater, & Williams, 2017) reported effects on parenting (ds .37-.70) but no such effect on child behaviour.
Minding the Baby® (Slade et al., 2005)	Pregnancy-2 years	Home-visiting programme targeting parental reflective functioning. Weekly-fortnightly visits delivered by pairs of social workers and nurses over two years.	In a two-year follow-up trial (Longhi et al., 2020) of 148 first-time mothers aged 14-25, no benefit of the intervention was found for maternal sensitivity, or cognitive development, but the intervention led to a reduction in child behaviour at age two, and sensitivity analysis revealed some evidence of improvement to security of attachment.

Intervention	Child age	Key components of delivery	Evidence base (with a focus on intervention delivery in the UK)
New Forest Parenting Programme (NFPP; Sonuga-Barke et al., 2006)	2-4 years	Involves 8-12 home-based visits, targeting parental limit setting, positive and constructive interactions, and motivation and scaffolding of attention and self-organisation.	In a small (Sonuga-Barke, Daley, Thompson, Laver-Bradbury, & Weeks, 2001; *N* = 78) trial, the programme was effective at reducing ADHD symptoms in three-year-olds and improving parents' sense of wellbeing but these effects were not replicated when delivered in primary care (Sonuga-Barke, Thompson, Daley, & Laver-Bradbury, 2004). A recent trial comparing NFPP, IY, and treatment as usual revealed modest effects on child behaviour (Sonuga-Barke et al., 2018).
Parent-infant psychotherapy (Fraiberg, 1980)	0-2 years	Dyadic intervention to target the parent-child relationship through the parents' internal working models. A psychotherapist observes and discusses the interactions.	In a small (Fonagy, Sleed, & Baradon, 2016; *N* = 76) trial, the 16-session intervention was found to promote parents' mental health and representations of their baby but not infant development. A review of delivery of PIP with parent-infant dyads (Barlow, Bennett, Midgley, Larkin, & Wei, 2015; *N* = 8 studies; 846 participants; mean child age = 24 months) found low-quality evidence for effects on secure attachment but no evidence for benefits in terms of other outcomes (maternal sensitivity and mental health, and infant behaviour and cognitive development).

1.5.4.1 Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline

Here I discuss the theoretical underpinnings and core components of delivery in the VIPP-SD programme.

1.5.4.1.1 Intervention's theoretical underpinnings

VIPP-SD is a manualised, interaction-focused programme that integrates attachment (Ainsworth et al., 1978; Bowlby, 1969/1982) and social learning and coercion (Patterson, 1982) theories into its messaging. It is designed to enhance parents' sensitivity and use of sensitive discipline (Juffer, Bakermans-Kranenburg, et al., 2017a). Programme themes for sensitive parenting focus on accurately observing and interpreting behaviour and emotions and responding in a sensitive way. The sensitive discipline component promotes a non-harsh but consistent response to challenging behaviour, with a focus on using inductive discipline and empathy.

1.5.4.1.2 Video feedback approach

VIPP-SD uses a video-feedback approach. Videotaped caregiver-child interactions are used by the therapist to provide feedback to the parent on their child's behaviour and their own parenting. Video footage can enhance observational insight, allowing therapists and caregivers to attend to even very subtle moments in the interaction. Video has been described as a 'looking glass' through which parents can observe emotions and behaviour with 'new' eyes (Juffer & Steele, 2014). The stop-start nature of the feedback allows positive moments to be clearly highlighted to the caregivers. Reinforcing these moments through video may increase the likelihood of caregivers continuing or increasing the use of their sensitive skills in day-to-day interactions with their child (Juffer et al., 2008b). In line with this, meta-analytic evidence indicates that interventions that use video feedback may be more effective in improving sensitive parenting than interventions without this method (Bakermans-Kranenburg et al., 2003; $d = 0.42$ vs. $d = 0.38$).

1.5.4.1.3 Six-session delivery

VIPP-SD is delivered across six sessions (four core, two booster) with standardised themes. 'More' does not necessarily translate to 'better' for psychological therapies. Engaging families in lengthy interventions can be resource-intensive, costly, and burdensome, decreasing the likelihood of parents completing the full intended dose (Juffer, Bakermans-

Kranenburg, et al., 2017a; O'Brien & Daley, 2011). Meta-analytic evidence suggests that short-term interventions (<16 sessions) with a narrow, behavioural focus may be more effective than lengthier programmes (Bakermans-Kranenburg et al., 2003; 0-16 session: d = 0.42, 16+ sessions: d = 0.21).

1.5.4.1.4 Home-based delivery

VIPP-SD is delivered in the home. This may allow for more natural and authentic day-to-day interactions to be captured on film (Juffer et al., 2008b). Home delivery may also provide a more comfortable, safer space for caregivers to receive personal feedback and practice new behaviours and observational skills (Juffer, Struis, Werner, & Bakermans-Kranenburg, 2017). This may increase the likelihood of caregivers integrating these behavioural changes into their daily interactions. Home-based delivery can also remove practical barriers for families and introduces a degree of flexibility (e.g., out of hours visits) that may be helpful for 'hard to reach' populations such as fathers (Panter-Brick et al., 2014).

1.5.4.1.5 Effectiveness of VIPP-SD

A meta-analysis of 12 randomised controlled trials testing VIPP-SD in 1,116 families revealed a combined effect size of d = .47 for improvements to sensitive parenting and d = .26 for reductions in behaviour problems (Juffer, Bakermans-Kranenburg, et al., 2017a). These effects appeared to be sustained over a longer-term follow-up, with a reported effect size of d = .25 for child outcomes. These findings provide promising evidence for the effectiveness of the VIPP model in improving parent and child functioning. However, it should be noted that all of these trials were undertaken in research settings, rather than public health contexts.

So far, I have considered the presentation and prevalence of behaviour problems in young children, and the important role parents can play in their children's development. However, the research discussed has focused exclusively on the thoughts, perceptions, and understandings of child development provided by adults (i.e., researchers, parents, teachers, clinicians). We know less about children's own perspectives of their development. The potential of children's participation in research forms the focus of the next section.

1.6 Integrating children's perspectives

Over the last few decades, there has been a rising commitment to include children themselves in matters that affect their lives. An emphasis within key political and policy arenas on the requirement and value of attaining children's perspectives almost certainly

bolstered these efforts (e.g., Children Act, 1989; European Commission, 2021; United Nations Committee on the Rights of the Child, 2005; United Nations Convention on the Rights of the Child, 1989). These legislative developments, paralleled with theoretical and methodological advances in our understanding, have led to an increased recognition of children's capabilities to enrich our understanding of their experiences (James, 2007). However, children's perspectives, and particularly those of very young children, are still seldom included in research, and rarely inform policy and practice decisions (Roberts, 2000).

The conceptualisation and study of children's perspectives is complex which may contribute to the hesitation of most mainstream researchers to engage with such approaches. Depicting these intricacies accurately and comprehensively is a challenge beyond the scope of this chapter. Instead, I centre my attention on two specific perspectives of children's participation: the sociology of childhood and developmental psychology. I then reflect upon some of the ethical and methodological challenges in this area and provide an overview of measures designed to elicit children's views, with a focus on play-based narrative assessment tools.

1.6.1 Why and how should we listen to children?

How and why children are invited to participate in research can influence the ways their contributions are interpreted and translated to construct meaning (James, 2007). Levels of participation range from adult-directed consultation, co-constructed participation, or independent self-initiation (Dockett, Einarsdottir, & Perry, 2009). Popular expressions like 'giving voice' or 'listening to' children can serve to simplify or mask several issues related to these processes. Most research on children's perspectives is "prompted, designed, analysed, and disseminated by adults" (Greene & Hill, 2005, p. 12). Rather than being passive, listening should therefore be viewed as an active process in which interpretation, meaning, and value is applied to what individuals offer (Rinalidi, 2006). The specific rationale and approach to including children in data gathering can be greatly influenced by the theoretical model researchers subscribe to.

Here, I consider the inclusion of children's voices from two (historically opposing) perspectives: the sociology of childhood and developmental psychology. Both schools of thought informed the present work. A central component of this thesis is the use of a doll's house play story stem measure with a sample of three- and four-year-olds. The measure was administered with the aim of gathering more information about children's behavioural

development. The information children provided using this play-based measure was coded using a quantitative framework and the data were compared to parent- and teacher-reports of child behaviour. In the origins of this research, the motivation to integrate children's perspectives into my studies was, to an extent, tied into core ideas related to childhood studies. However, the research's execution, approach, aims, and analyses became more closely aligned to the views and practices more typically seen in developmental psychology. Each theoretical perspective is now considered in turn.

1.6.1.1 The (new) sociology of childhood and children's participation

The new social studies of childhood regard children as experts of their own lives, who are knowledgeable individuals, capable of expressing unique perspectives, and making valuable contributions to society (Einarsdóttir, 2007; James, 2007; James, Jenks, & Prout, 1998; Prout, 2011; Tay-Lim & Lim, 2013). This perspective argues that children's experiences are important to study in their own right, rather than solely as a way to enhance our understanding of children as developing adults, in order to promote 'optimal outcomes' (Tatlow-Golden & Montgomery, 2021). Theorists have focused less on children as 'becomings' (i.e., adults in the making) and more on children as 'beings' (Qvortrup, 2009; Tangen, 2008), although this conceptualisation has been critiqued for portraying a completeness and stability to 'being' (Tatlow-Golden & Montgomery, 2021). Research practices that are commonplace in developmental psychology, including quantitative approaches and statistical analysis, have received criticism under this perspective for failing to capture the richness and diversity of individual lives (James, 2007). The sociological perspective also strives to account for the ways childhood is shaped by politics and policy, recognising the social, historical, economic, cultural, and moral structures underlying these experiences (Quennerstedt & Quennerstedt, 2014).

The studies in this thesis stem from an understanding that children can provide valuable and meaningful information when afforded authentic opportunities to do so. Their unique insights are considered to have the potential to enhance our understandings that stem from adults' viewpoints. However, in this thesis, the perspectives of parents and teachers are also regarded as highly important and wider features of the sociocultural approach are not accounted for. I view the collection, analysis, and interpretation of children's data to relate more closely to the practices of children's participation in developmental psychology studies, an outline of which is now provided.

1.6.1.2 Developmental psychology and children's participation

Embedding children's perspectives into investigations in developmental psychology research has been widely viewed as a pathway to enhance our understanding of their development. Children's perspectives are often considered in the context of child-related outcomes and variables, rather than through a detailed examination of individual children's experiences. Nilsson et al. (2015) suggest that acquiring children's views through instruments that rate their behaviour and feelings, as is the case in the present thesis, allows us to 'have' a child's perspective but not to 'take' a child's perspective, something that is achieved only when children can express themselves without boundaries.

In psychology, children's development is mostly understood through the adults around them including parents and caregivers, teachers, and clinicians (Luby, Belden, Sullivan, & Spitznagel, 2007). Each informant's perspective offers a partial view of the child's behaviour related to the situational specificity of the informant-child relationship (Arseneault, Kim-Cohen, Taylor, Caspi, & Moffitt, 2005). Children's perspectives can strengthen multi-informant approaches, particularly when we consider: the ways parents' perceptions can be biased by their own experiences (e.g., parental psychopathology; Ringoot et al., 2015); the typically low response rate of teachers (Arseneault et al., 2005); and the modest degree of convergence across informants (Achenbach, McConaughy, & Howell, 1987; De Los Reyes & Kazdin, 2005).

In this thesis, the incorporation of children's perspectives involved translating children's words into numbers, with conclusions drawn from quantifiable data. Through these processes, the richness and diversity of children's responses were decontextualised. However, children's perspectives rarely make it to the point of informing policy and practice. Quantitative data tends to be more readily prized by policymakers and advocates looking to communicate a clear message and justify decision making (Kleine, Pearson, & Poveda, 2016). Quantitative approaches may offer a unique opportunity to bring children's insights to the attention of those who do not ordinarily recognise their value.

In sum, the current research is underpinned by features of multiple conceptual frameworks related to listening to and respecting the value of children's voices. Next, I consider some of the ethical challenges children's research participation can bring.

1.6.2 What are some of the key ethical considerations when integrating children's perspectives?

The participation of young children can bring a unique set of ethical tensions. A common theme within these ethical considerations is how to avoid children's participation being a tokenistic aspect of the research (Dockett et al., 2009). Key issues relate to: researchers attaining consent/assent for children's participation; acknowledging the (un)representativeness of participating samples of children; challenging the (limitations of) adult-directed methods and analysis; understanding the individual and societal impacts of children's participation; and ensuring that the completeness of findings are not overstated (i.e., it is impossible for children to express all aspects of their experiences, particularly when reliant on researchers' own interpretation and understanding) (Dockett et al., 2009; Johansson, 2011). These issues are not always easily solvable or avoidable. Instead, researchers are encouraged to engage in a process of ongoing reflexivity to recognise and engage in these questions and remain mindful of their own influence in the research (Dockett et al., 2009) How does an adult-interpretation of data change children's contributions? To what extent do children understand why they are providing information and how it will be used? Has it been recognised that some children cannot speak for all children? Are there opportunities embedded in the research process that allow children themselves to interpret the data? What can be done to amplify the voices of overlooked groups of children?

The research undertaken as part of this thesis aligned with the ethical guidelines set out by the governing bodies in health research. However, the nuance of aspects of the ethical considerations discussed in this section only became clearer to me after I had completed data collection with the children. Thus, through this thesis, I make efforts to acknowledge where processes could have been enhanced to strengthen the ethical integrity of the research.

1.6.3 What are some of the methodological challenges of integrating children's perspectives?

Adults have a tendency to underestimate the cognitive and communicative capabilities of very young children (Greene & Hill, 2005). The view of children as competent and valuable contributors to research was discussed earlier in this chapter. However, using research methods designed and developed for adults would not allow these proficiencies to be recognised, particularly in younger or non/minimally-verbal children (Stephenson, 2009). To facilitate their participation, researchers must adopt developmentally sensitive methods of

data collection and interpretation that maximise children's opportunities to express their views and expertise. "The immaturity of children is a biological fact of life, but the ways in which this immaturity is understood and made meaningful is a fact of culture" (Prout & James, 2015, p.6). These methods must account for young children's competencies and differences in their attentional, cognitive, and linguistic capabilities. Methods must also be engaging. In contrast to adults, who may be willing to sit through dull and laborious measures, children are unlikely to be motivated to complete measures simply for the good of research. A selection of qualitative and quantitative methods that have been adapted to suit the developmental capabilities and interests of young children are considered next.

1.6.4 How can young children's perspectives be elicited?

Measure selection is influenced by the research aims and questions, the time and resource available, the characteristics and needs of participating children, and the framework adopted by the research team (Christensen & Prout, 2002). A common feature is that methods used to engage very young children are rarely solely reliant on verbal proficiency, are often creative in nature, and generally capitalise on children's penchant for playful activities. These approaches can allow even very young children to articulate their perspectives with relative ease, even when communicating abstract concepts (Tay-Lim & Lim, 2013). Table 2 presents some examples of methodologies used with young children, with an exemplar research case offered for each approach. Next, I focus closely on just one methodological approach: the elicitation of narrative representations.

Table 2

Methodological Approaches to Elicit the Perspectives of Young Children with Exemplar Research Cases

Type	Brief description	Research example
Drawings	Inviting children to 'draw and talk' about a topic can be a particularly helpful way to facilitate conversations with children about concepts they might otherwise find difficult to discuss – allowing them to express themselves without language. Drawings are commonly viewed as representational tools to elicit children's meaning and perceptions.	Dockett and Perry (2005) invited 39 kindergarten children to draw something about school, to consider children's experiences and expectations in their first year of school. They found that children "represent school in different ways, and drawing provides a forum where individual experiences of school can be reflected and acknowledged" (p. 13).
Photography	Photography can provide insight into the world through the child's eyes by allowing children to capture images of things and experiences that are important to them. The photographs can then offer a 'talking point' for the researcher and the child in discussions about the concepts of interest.	Pyle (2013) invited 32 kindergarten-aged children to take photos of the important places, activities and people in their classrooms. These photos were then used in interviews with the children to find out more about the kindergarten classroom environment. It was concluded that the combination of photos and interviews offered a way for children to influence and contextualise the data collected.
Pictorial stimuli	Pictures can be used during interviews to stimulate children's responses. Researchers can deliver standardised prompts while exploring the pictures or allow the child to respond freely.	Harrison and Murray (2015) used a series of illustrations that depicted typical classroom events in interviews with 101 children aged four to six years. They concluded that the tool offered a way to gather children's perceptions of their own wellbeing at school.
Puppet interviews	In interviews, two puppets can be used to allow children to complete measures in an interactive way (e.g., Puppet 1: 'I like ice cream', Puppet 2: 'I don't like ice cream. How about you?'). This method is engaging and less reliant on children's verbal skills, as they can just point to the puppets.	Ringoot et al. (2017) used a puppet interview with six-year-olds to assess behaviour problems. Child-reported problems were associated with caregiver-reported clinical difficulties and predicted treatment referral two years later. It was concluded that systematic play-based interviewing could enhance child assessments.
Storygames	Children are invited to each give a line of a story, going from child to child until it reaches completion. Such approaches are seen as a helpful alternative to asking children for 'your story', something that can be inappropriate for ethical reasons.	Veale (2005) used storygames with children aged 7-12 years in Rwanda. Stories told related to their experiences of genocide. It was speculated that the sense of the 'game' may have been protective for children as their stories were constructed without distress.

1.6.4.1 How can narrative-based representational approaches tell us more about children's development?

The use of representational, narrative-based measures forms a core component of this thesis and will be considered in close detail in subsequent chapters. A brief overview of some of the procedural and theoretical considerations for their use are provided here.

1.6.4.1.1 What do narrative approaches look like?

Narrative approaches cater for several aspects of children's attentional, linguistic, and cognitive abilities, facilitating communication through the use of play, storytelling, and engaging stimuli (e.g., pictures, puppets, dolls) (Goldstein, 2012; Murray, Woolgar, Briers, & Hipwell, 1999). Dolls play techniques involve encouraging children to manipulate figures while telling a story. Analysis of these narratives is considered to provide insight into how children make sense of their world (Buchsbaum & Emde, 1990). Children develop the skills for narrative production by about three years of age making these methods suitable for use with preschool-aged populations (Bretherton & Beeghly, 1982). The 'story stem' approach involves the delivery of an emotionally evocative story beginning and an invitation to the child to show and tell what happens next. An empirical evidence base supporting the use of these tools has formed over the last few decades (see Yuval-Adler & Oppenheim, 2015) and will be considered further throughout this thesis.

1.6.4.1.2 What are some of the theoretical underpinnings of narrative approaches?

Theoretically, the delivery of story beginnings are believed to act as a catalyst for the child to draw upon their internal knowledge and understanding of similar situations. The precise ways in which these 'inner worlds' are described and conceptualised differs across the field and will be considered in finer detail in subsequent chapters. In brief, it is proposed that children's previous interactions, experiences, and relationships are internalised to form an inner system that guides children's approach, interpretation, and response to their external worlds (Ainsworth et al., 1978; Bowlby, 1969/1982; Bretherton, Ridgeway, & Cassidy, 1990; Erickson, Sroufe, & Egeland, 1985). The presentation of the 'dilemma' in the story stem approach is considered to activate aspects of these internal representations that reveal schemas underlying the child's typical behaviours (Murray et al., 1999).

Overall, play-based narrative techniques offer a developmentally sensitive way to gain insight into the ways in which very young children understand and interpret their worlds. However, the use of these methods has mostly been restricted to low-risk, school-aged

populations of small sample size. Less is known about the use of these measures with preschool-aged children, especially those considered to be at risk of behavioural difficulties. Obtaining insight through these measures during a particularly sensitive time in development may enhance our understanding of key factors involved in child development. As well as this, integrating these methods into a large-scale study of a parenting intervention could offer a unique opportunity to consider the ways in which (changes to) children's external environments may be associated with their internal functioning.

1.7 The current thesis

The research forming this thesis is guided by five broad questions:

1. How can children's perspectives be integrated into large-scale developmental psychology research?
2. What can children's perspectives tell us about their perceptions of their own behaviour?
3. How do children's perspectives relate to their early experiences of parenting?
4. Can a brief video-feedback intervention influence aspects of children's narrative representations?
5. To what extent do changes to parental sensitivity explain intervention effects on children's external and internal functioning?

The subsequent chapters aim to address elements of these questions. Chapter 2 provides an overview of the Healthy Start, Happy Start study, the randomised controlled trial in which the samples for the studies within this thesis were drawn. Chapter 3 introduces an adapted version of a story stem assessment battery used with a sample of three- and four-year-olds, examining links between constructs derived from children's narratives and sociodemographic factors, as well as parent- and teacher-reported data. Chapter 4 examines longitudinal associations between early child behaviour, parental discipline styles, and children's narrative representations, as well as the effects of a parenting intervention on constructs derived from children's data. Chapter 5 builds on these findings and examines the mediating role of parental sensitivity on children's internal and external functioning. Chapter 6 offers a synthesis of key findings of studies and a discussion of the strengths and limitations, and implications for future research, policy, and practice of this thesis.

CHAPTER 2: Healthy Start, Happy Start: An overview

2.1 Chapter rationale

The aim of this chapter is to provide a comprehensive overview of the Healthy Start, Happy Start study, a large-scale, pragmatic, randomised controlled trial. The findings of this thesis are drawn from this sample and thus there is value in starting with a detailed consideration of the trial's design and methodology.

2.2 Introduction

In this chapter, I provide an overview of the Healthy Start, Happy Start (HSHS) trial. The studies presented in this thesis were embedded within the broader HSHS study. More detailed methodological information will be provided where most relevant in subsequent chapters. There may be similarities in descriptions of the trial's methodology given in this chapter and other study-related publications that I have co-authored (O'Farrelly, Barker, et al., 2021; O'Farrelly, Watt, et al., 2021; Rao, Barker, O'Farrelly, & Ramchandani, 2021). Figure 2 outlines my contributions to the HSHS study, demarcated by my time as a Research Assistant (Feb 2016-Oct 2017) and then, predominantly, as a doctoral candidate (Nov 2017-Jul 2021).

The rationale for conducting the wider HSHS trial was rooted in the need for effective interventions for children's early behaviour problems. As discussed in the previous chapter, behaviour problems are common and can have enduring consequences for children, their families, and society (Sadler et al., 2018). However, the UK is lacking in effective supports that can be delivered at scale in public health services at the earliest signs of these difficulties. The primary aim of HSHS was to test the clinical- and cost-effectiveness of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2008a; Juffer et al., 2008b). The programme aims to promote parental sensitivity and a sensitive discipline style. In HSHS, it was hypothesised that receipt of the programme would reduce children's behaviour problems.

I begin this chapter by outlining the study's design, ethics, and governance. I then provide an overview of the eligibility, recruitment, retention, consent, and demographic characteristics of participating families, before offering details of research measures and data management and handling.

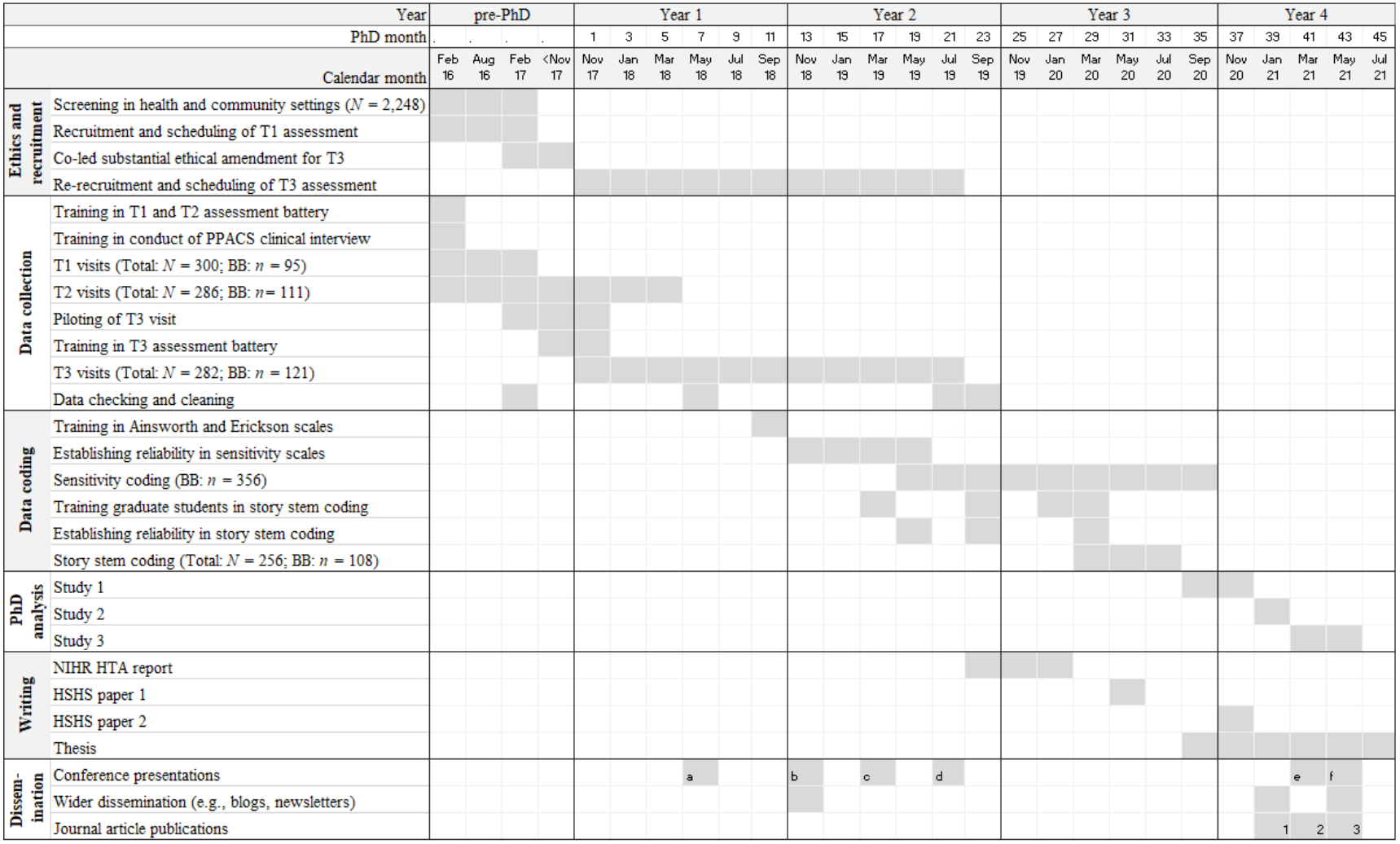


Figure 2. An overview of my contributions to the HSHS trial and my embedded PhD research.

Note: Where appropriate, *N*s are broken down into total number of cases and my own contributions (BB)
a = 16th WAIMH World Congress; b = Emmanuel Miller ACAMH; c = SRCD Biennial meeting; d = VIPP symposium; e = SRCD Biennial meeting; f = Play symposium
1 = O'Farrelly, Watt, et al. (2021); 2 = Rao, Barker, et al. (2021); 3 = O'Farrelly, Barker, et al. (2021)

2.3 Method

2.3.1 Study design

The Healthy Start, Happy Start study (HSHS; see Ramchandani et al., 2017 for trial protocol) was a pragmatic, researcher-blind, multi-site, two-arm parallel group randomised controlled trial (RCT). Its primary aim was to test the effectiveness of a brief video-feedback intervention, with caregivers of young children aged 12 to 36 months who were considered to be at an increased likelihood of developing enduring behaviour problems.

2.3.2 Ethics and trial governance

The trial was funded by the National Institute for Health Research, Health Technology Assessment (NIHR HTA; ref: 13/04/33) and sponsored by Imperial College London. Ethical approval for the trial was granted by Riverside Research Ethics Committee (REC ref: 14/LO/2071; Appendix A). I co-led on the submission of a substantial ethical amendment that allowed for the inclusion of exploratory outcome measures used in the analyses throughout this thesis. The trial was registered with the ISRCTN clinical trial registry (ISRCTN58327365) with individual approvals for participating sites provided by the NHS Health Research Authority.

2.3.3 Eligibility criteria

Participants were children aged 12 to 36 months with elevated parental reports of externalising behaviour problems at screening, and their caregiver(s). In a portion of families, there were two caregivers willing and available to participate (n = 51; 17%), while the remainder participated with one caregiver (n = 249; 83%). All participating families met the following eligibility criteria:

2.3.3.1 Inclusion criteria

1. Participating caregiver(s) aged ≥ 18 years.
2. Child aged between 12 and 36 months at screening.
3. Child scored in the top 20% of population norms on the externalising subscale of the parent-reported Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997).
4. Written informed consent provided by participating caregiver(s).

2.3.3.2 Exclusion criteria

1. Child or caregiver had severe sensory impairment, learning disability, or language limitation, which was sufficient to preclude participation in the trial.
2. Child's sibling was already participating in the trial.
3. Family was participating in active family court proceedings.
4. Caregiver was participating in another closely-related research trial and/or was currently receiving an individual video-feedback-based intervention.

2.3.4 Participant recruitment

Recruitment for the study was conducted within six National Health Service (NHS) trusts; Central and North West London NHS Foundation Trust, Whittington Health NHS Trust, North East London NHS Foundation Trust, Oxford Health NHS Foundation Trust, Cambridgeshire and Peterborough NHS Foundation Trust, and Hertfordshire Community NHS Trust. Recruitment efforts were focused across seven areas within these trusts: The London Boroughs of Camden, Hillingdon, Islington, and Barking and Dagenham, alongside Oxfordshire, Peterborough, and Hertfordshire. Recruitment occurred over two phases: the screening phase and the trial phase.

2.3.4.1 Phase one recruitment: Screening phase

In the screening phase, families were approached at health visiting developmental review sessions. Families typically attend two developmental reviews during toddlerhood (at ~10-12 months and 24-27 months of age). Caregivers were either asked by their health visitor to complete a screening pack as part of the review itself or were approached in the waiting room by a member of the research team. The screening pack contained the phase one participant information sheet, consent form, and a screening questionnaire (SDQ; Goodman, 1997). In four participating sites, local health visiting services also used mailshots to post a letter to caregivers of children aged 12 to 36 months inviting them to complete the HSHS screening. Face-to-face recruitment was also completed across other NHS and community services, including children's centres and libraries, and study information was shared through local social media platforms.

2.3.4.2 Phase two recruitment: Trial phase

In the trial phase, families whose children scored in the top 20% of population norms on the externalising problems subscale of the SDQ at screening were followed up by a

member of the research team via telephone. During this phone call, the researcher established whether the family met the full eligibility criteria for participation in phase two of the study. Eligible families who were willing to participate in the trial were sent more information about the study via email or post ahead of their scheduled baseline assessment. The study recruited to target (N = 300). Figure 3 provides an overview of participant recruitment and retention during the study, in accordance with CONSORT guidelines. Most participants were recruited through health visiting services (30% at developmental reviews and 25% through service mailshots) and children's centres (30%) (see Table 3). Most families lived in London (75%) with smaller numbers in Oxfordshire, Hertfordshire, and Peterborough (see Table 4).

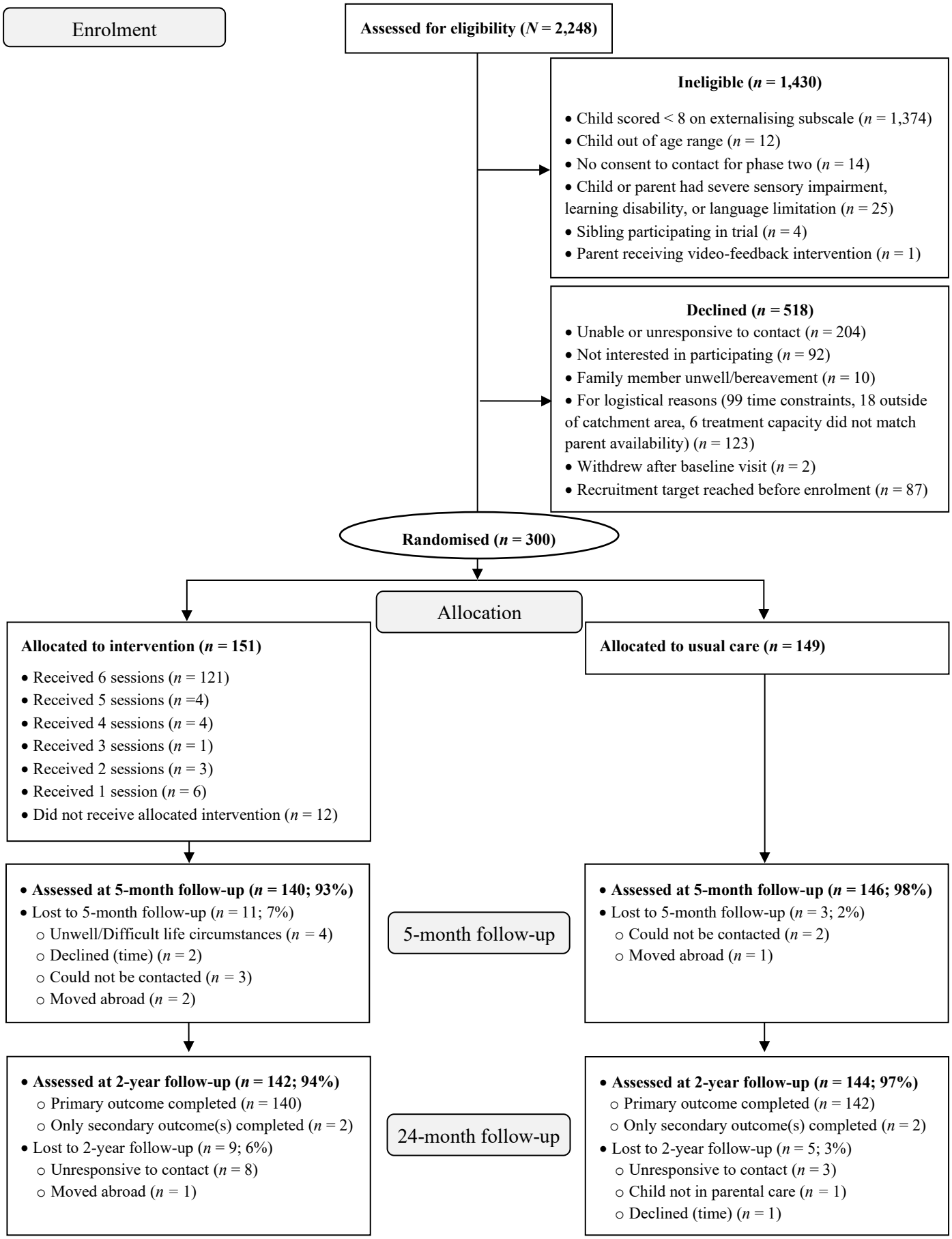


Figure 3. Flow of all HSHS participants through trial.

Table 3

Recruitment Pathway for Participants Recruited into the HSHS Trial

Recruitment route at screening	Number of families, *n* (%)
Routine developmental review by health visitors	91 (30)
Health visitor mailshot	74 (25)
Children's centre	91 (30)
Other clinic/community venues	23 (8)
Online advert	14 (5)
Word of mouth	3 (1)
Other	4 (1)

N = 300

Table 4

Recruitment Location of Participating HSHS Families

Recruitment site (NHS trust)	Number of families, *n* (%)
Barking and Dagenham (North East London NHS Foundation Trust)	14 (5)
Camden (Central North West London Foundation Trust)	83 (28)
Hertfordshire (Hertfordshire Community NHS Trust)	7 (2)
Hillingdon (Central North West London NHS Foundation Trust)	22 (7)
Islington (Whittington Health NHS Trust)	105 (35)
Oxfordshire (Oxford Health NHS Foundation Trust)	43 (14)
Peterborough (Cambridgeshire and Peterborough NHS Foundation Trust)	26 (9)

N = 300

2.3.5 Sample characteristics

The sociodemographic data of the 300 participating families are presented in Table 5. Here, I consider the representativeness of this sample compared to the wider UK population. See Appendix B for the sociodemographic data of children's second caregivers.

Child age and gender. At baseline, children were 12 to 36 months old (M = 23.0 months; SD = 6.7 months). There were slightly more male than female children participating (54% male), possibly due to the fact behaviour problems are more common in boys. A recent study indicated that this higher prevalence is detectable in children as young as two to four years, with a higher incidence of both behavioural (boys: 3.1%; girls: 1.9%) and hyperactivity disorders (boys: 0.6%; girls: 0.3%) seen in male children (Sadler et al., 2018).

Child ethnicity. Children's ethnicity was reported by their caregiver as follows: 6% Asian, 6% Black, 20% Mixed heritage ethnicity, 65% white, and 3% Other. Data released on the registered birth characteristics of all children born in the UK between 2014 and 2018 is used as a comparison as these closely correspond with the birth years of the HSHS children (2013-2016). Between 2014 and 2018 babies' ethnicities in the UK were recorded as follows: 9% Asian, 4% Black, 72% white, and 11% Other (Office for National Statistics, 2020b). The HSHS sample is more ethnically diverse than the general UK population possibly reflecting the increased diversity in London (19% Asian; 13% Black; 5% Mixed heritage; 60% white; 3% Other; Census, 2011a), where most participating families lived.

Caregiver relationship status, education, and employment. Participating caregivers were mostly female (96%) and the child's biological or adoptive mother (96%) or father (4%). More HSHS caregivers were married or living with a partner (85%) than is seen in the general population (78%) (Office for National Statistics, 2017). However, the population data are based on families with a child aged 0 to 17 years. Relationship breakdown between coparents is more common as children get older and so we would expect higher rates of cohabitation in the toddler years. As is common in research studies, a higher percentage of this sample held a graduate-level qualification (64%) compared to the general population of a similar age (41%; Census, 2011b). In the HSHS sample, 59% of primary caregivers were employed (including being on paid parental leave) while 37% reported looking after the home and family. This is a slightly lower employment rate than that seen in the general population, where 67% of women with at least one child aged 0 to 2 years were employed, while 30% were looking after the family home (Office for National Statistics, 2019).

Table 5

Sample Characteristics of Participating Children and their Primary Caregivers

Characteristic	Trial arm		
	VIPP-SD (n = 151)	Usual care (n = 149)	All (N = 300)
Child characteristic			
Sex (female), *n* (%)	75 (50)	62 (42)	137 (46)
Age at baseline (months), mean (SD)	22.8 (6.8)	23.2 (6.5)	23.0 (6.7)
Ethnicity, *n* (%)			
Asian	9 (6)	8 (5)	17 (6)
Black	3 (2)	15 (10)	18 (6)
Mixed	36 (24)	25 (17)	61 (20)
White	100 (66)	94 (63)	194 (65)
Other	3 (2)	7 (5)	10 (3)
Primary caregiver characteristic			
Sex (female), *n* (%)	143 (95)	144 (97)	287 (96)
Age at baseline (years), mean (SD)	33.7 (5.6)	34.7 (5.9)	34.2 (5.8)
Parental status, *n* (%)			
Parent (including step or adoptive)	151 (100)	149 (100)	300 (100)
Ethnicity, *n* (%)			
Asian	15 (10)	16 (11)	31 (10)
Black	3 (2)	15 (10)	18 (6)
Mixed	11 (7)	11 (7)	22 (8)
White	114 (75)	103 (69)	217 (72)
Other	8 (5)	4 (3)	12 (4)
Relationship status, *n* (%)			
Married/civil partnership/cohabiting	128 (84)	127 (85)	255 (85)
Single	12 (8)	17 (11)	29 (10)
In relationship but not cohabiting	10 (7)	1 (1)	11 (3)
Divorced/widowed/legally separated	1 (1)	4 (3)	5 (2)

Characteristic	Trial arm		
	VIPP-SD (n = 151)	Usual care (n = 149)	All (N = 300)
Employment status, *n* (%)			
Employed	86 (57)	76 (51)	162 (54)
Looking after home and children	56 (37)	56 (37)	112 (37)
Paid parental leave	6 (4)	10 (7)	16 (5)
Full-time student	3 (2)	7 (5)	10 (4)
Highest qualification, *n* (%)			
Graduate	92 (61)	99 (67)	191 (64)
A-level/NVQ/BTEC	42 (28)	36 (24)	78 (26)
GCSE or lower	17 (11)	14 (9)	31 (10)

2.3.6 Informed consent

Participating caregivers provided written informed consent on behalf of them and their child for each phase of the study. All participant materials were ethically approved. In the screening phase, families were provided with a phase one participant information sheet and a consent form (see Appendix C). Families were able to ask questions in-person if recruited face-to-face or were provided with the study team's contact details if recruited through online pathways. In the trial stage, the phase two participant information sheet and consent form (see Appendix D) were explained to the participant at the beginning of their baseline assessment. Families were able to ask questions and were reminded of their right to withdraw at any stage. The limits of confidentiality were also explained to participants. Researchers collected information about the child's behaviour, the parent-child relationship, caregivers' approach to discipline, and parental mood. Clinically significant levels of distress in caregivers or safeguarding risks to the child could become apparent through elevated measure scores or through general conversation. Standardised operating procedures were in place to ensure researchers took the necessary steps to support families. All concerns were discussed with the trial manager or principal investigator and referrals were made to participants' healthcare services as appropriate. In addition, all participating families received localised debrief sheets outlining family and mental health support in their area.

2.3.7 Baseline and follow-up assessments

Baseline assessment visits were conducted between July 2015 and July 2017 and follow-up data was collected between December 2015 and July 2019. Research assessments were mostly conducted in families' homes and took between two and two-and-a-half hours to complete.

2.3.8 Measures

Table 6 outlines the timing of the administration of trial measures relevant to the current thesis. This section focuses on six measures used across the studies presented in this thesis. These measures are described again in subsequent chapters but, for brevity, in less detail. The story stem assessment battery is a central component of this thesis and forms the focus of the next chapter. Because of this, it is first introduced in the following chapter rather than here.

2.3.8.1 Measures of child behaviour

2.3.8.1.1 Questionnaire measure 1: Child Behaviour Checklist

The Child Behaviour Checklist (CBCL/1.5-5; Achenbach & Rescorla, 2000, 2001) is a 100-item measure of social, emotional, and behavioural problems. Respondents are asked to rate how true the behaviour is of the child in the last two months on a three-point scale (0 = not true, 1 = somewhat or sometimes true, 2 = very true or often true). Total, externalising, and internalising scores are generated. A key strength of this measure is that it has been previously utilised with children as young as 12 months of age (van Zeijl et al., 2006). It is well-validated, with good test-retest reliability (mean: r = .85), moderate cross-informant correlations (r = .61), and acceptable score stability over 12 months (r = .61) (Achenbach & Rescorla, 2000). The validity and consistency of the preschool version has also been examined across two large-scale studies of over 19,000 children aged 18 to 71 months. These studies reported good internal consistency (Cronbach's α = .84-.94) and confirmatory factor analysis indicated an acceptable to good model fit (Ivanova et al., 2010; Rescorla et al., 2011). Further study of the suitability of this measure in children as young as 12 months old would be beneficial to enhance our understanding of its use with very young populations

2.3.8.1.2 Questionnaire measure 2: Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item measure of behaviour problems. Respondents rate the child's behaviour over the last six months using a three-point scale (0 = not true, 1 = somewhat true, or 2 = certainly true). Five

subscale scores are generated: conduct problems, emotional symptoms, hyperactivity, peer problems, and prosocial behaviour. Key strengths of the SDQ are its brevity and its use of positive phrasing on items (Goodman & Scott, 1999). The 4-17-year-old version of the SDQ has good psychometric properties (Goodman, 2001; Lundh, Wångby-Lundh, & Bjärehed, 2008; Muris, Meesters, & van den Berg, 2003; Yao et al., 2009). The current study used the 2-4-year-old version, which features adjustments to three items made by the measure's developer to improve the developmental suitability of items (e.g., 'Often lies or cheats' changed to 'Often argumentative with adults'). Comparative to the 4-17-year-old version, studies exploring the psychometric properties of the 2-4-year-old version are still somewhat limited. However, findings from those studies that have been conducted provide an indication for its suitability for preschool-aged children. The original five-factor structure has been replicated across studies (Croft, Stride, Maughan, & Rowe, 2015; Ezpeleta, Granero, de la Osa, Penelo, & Domènech, 2013; Klein, Otto, Fuchs, Zenger, & von Klitzing, 2013) and the internal reliability for subscales has been found to be acceptable (α range = 0.58–0.84; D'Souza, Waldie, Peterson, Underwood, & Morton, 2017; Klein et al., 2013). A cohort study of 16,659 children revealed strong positive correlations between SDQ scores at three and five years (conduct r = .73; hyperactivity r = .71) that were comparable to those seen between reports at five and seven years (conduct r = .80; hyperactivity r = .79), providing support for the measures predictive validity (Croft et al., 2015). Conduct problems and hyperactivity subscales measured at three years old also independently predicted developmental and clinical outcomes two years later. Further study of the suitability of this measure in children as young as one and two years old would be beneficial to enhance our understanding of its psychometric properties when used in toddler-aged populations.

2.3.8.1.3 Interview measure: Preschool Parental Account of Childhood Symptoms

Children's externalising behaviour problems were also assessed using a semi-structured, researcher-led clinical interview conducted with children's primary caregivers. The Preschool Parental Account of Childhood Symptoms (PPACS; Sonuga-Barke, Lamparelli, Stevenson, Thompson, & Henry, 1994; Taylor, Sandberg, Thorley, & Giles, 1991) has been found to have high inter-rater reliability (*rs* .92-.98), test-retest reliability (*r*s .62-.78), and internal consistency (Cronbach's α 0.87-.89) (Sonuga-Barke et al., 1994; Taylor et al., 1986b; Taylor et al., 1986a). Moderate associations were found between subscale scores of hyperactivity (r = .57) and conduct problems (r = .49) and related questionnaires (Sonuga-Barke et al., 1994). The measure was found to identify children with hyperkinetic disorder

with an estimated 0.96 specificity and 0.85 accuracy, and children with conduct disorder with a specificity of 0.71 and accuracy of 0.87 (Giannopoulou, 2003; Taylor et al., 1991).

For the HSHS study, modifications were made through piloting and discussion to form the PPACS Infant and Toddler version. Questions related to behaviours not often seen in very young children (e.g., telling lies and stealing) were removed. The reference timeframe was altered from the previous six months to the previous four months to accommodate for the rapid rate of behavioural development in toddlerhood. Behavioural examples used to guide scoring were revised to reflect developmental expectations in younger children (e.g., refusals to 'tidy up toys' and 'put on an item of clothing' were added as examples of disobedience).

The first half of the interview focused on inattention, hyperactivity, and fidgeting when the child was watching television, playing alone and with siblings/peers, at mealtimes, and when sleeping. The second half of the interview assessed the frequency and severity of conduct-related problems, exploring behaviours related to temper tantrums, defiance, destructiveness, aggressiveness, and problems at bedtime. For each question, the child's primary caregiver was asked to describe a concrete example of the child's behaviour within the last week. This was designed to act as an anchor for the parents' recall. Caregivers were then asked to report on the representativeness of this behaviour over the last four months to ensure the example was characteristic of the child. This approach meant scores were not reliant on parents' global impressions or clinical judgement of their children's behaviour. Instead, interviewers used their training to rate the described behaviour according to operationalised criteria with written descriptions of behaviours at varying thresholds of clinical difficulty. As an example, for temper tantrums, the severity scale consisted of: 0 = none, 1 = mild (e.g., shouting, waving arms), 2 = moderately severe (e.g., throwing things, kicking), and 3 = severe (e.g. breaking things, aggressive behaviours), while the frequency was scored on a scale of 0 = none or less than once a week, 1 = on 1 or 2 days per week, 2 = on 3 to 6 days a week, 3 = daily, and 4 = more than once a day. Interviews generally took 45 to 60 minutes to complete. Five researchers conducted interviews during the trial after undertaking two days of training and three supervised cases. All interviews were audio-recorded. Good to excellent inter-rater reliability was achieved at each time point (intra-class correlation coefficients for baseline: .93-.97; 5-month follow-up: .95-.96; and 24-month follow-up: .72-.98), measured using a randomly selected set of interviews (10%; $N = 30$ per time point).

2.3.8.2 Parenting practices: Parenting Scale

The Parenting Scale (Arnold, O'Leary, Wolff, & Acker, 1993) is a self-reported measure of parental discipline style consisting of 30 items on seven-point Likert scales. For each item, a situation (e.g. When my child misbehaves…) is anchored by both an 'effective' (e.g. I speak to my child calmly) and 'ineffective' (e.g. I raise my voice or yell) approach to discipline at each end of the scale. A total score and three subscale scores for laxness (permissive discipline), over reactivity (harsh discipline), and verbosity (overly wordy responses to misbehaviour) are produced. The measure has been found to have acceptable test-retest reliability and internal consistency (Arnold et al., 1993; Prinzie, Onghena, & Hellinckx, 2007), stability over six months ($rs \geq .63$, $ps < .001$; Lorber, Xu, Slep, Bulling, & O'Leary, 2014), and has been associated with measures of child behaviour and parental wellbeing (Prinzie et al., 2007; Lorber et al., 2014). The subscales were originally validated against home-based observations of parent-child discipline situations with children aged 18-48 months (Arnold et al., 1993). Subsequent studies supported a two-factor structure of Lax and Overreactive parenting but did not replicate the Verbosity factor found in the original development study (Harvey, Danforth, Ulaszek, & Eberhardt, 2001; Reitman et al., 2001; Steele, Nesbitt-Daly, Daniel, & Forehand, 2005). It has been proposed that this factor may only be relevant for children under three years of age (Rhoades & O'Leary, 2007).

2.3.8.3 Parental depression: Patient Health Questionnaire-9

Self-reported depression severity and symptomatology was measured by the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001). Caregivers were asked how often during the last fortnight they were bothered by nine symptoms, scored from 0 (not at all) to 3 (nearly every day). The nine items correspond with the criterion for depression outlined in the Diagnostic and Statistical Manual of Mental Disorders IV (4th ed.; DSM-IV; American Psychiatric Association, 2004). This brief questionnaire has been found to have good criterion and concurrent validity, and excellent internal reliability (Cronbach's $\alpha = 0.86$-0.89) and test-retest reliability ($r = 0.84$) (Spitzer, Kroenke, Williams, & Patient Health Questionnaire Primary Care Study Group, 1999; Spitzer et al., 2000).

2.3.8.4 Parental anxiety: Generalised Anxiety Disorder-7

Self-reported anxiety severity and symptomatology was measured by the Generalised Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006). Items reflect the symptom criteria outlined in the DSM-IV (American Psychiatric Association, 2004).

Respondents were asked how often they had experienced the symptoms in the last two weeks, ranging from 0 (not at all) to 3 (nearly every day). The measure has good to excellent internal consistency (Cronbach's α = .92), test-retest reliability (ICC = 0.83) and convergent validity (*rs* .72-.74) (Spitzer et al., 2006).

Table 6

Timing of Data Collection for Variables Analysed in this Thesis

Broad construct	Variable	Measure	Time point		
			BL	5M	24M
Demographic information	Caregiver and child sociodemographic information	Interview	X		
Child behaviour	Behaviour problems	CBCL[a]	X	X	X
	Behaviour problems	PPACS[b]	X		X
	Behaviour problems and prosociality	SDQ[c]	X		X
	Observed behavioural dysregulation	Story stems[d]			X
Parenting	Dysfunctional parental discipline practices	Parenting scale	X		X
	Parental sensitivity	Erickson scales	X	X	
Caregiver wellbeing	Symptoms of depression	PHQ-9[e]	X		X
	Symptoms of anxiety	GAD-7[f]			X
Children's representations	Child representations of behaviour, prosociality and narrative coherence	Story stems[d]			X

Note. BL = baseline assessment, 5M = five-month follow-up, 24M = two-year follow-up

a CBCL = Child Behavior Checklist; b PPACS = Preschool Parental Account of Childhood Symptoms; c SDQ = Strengths and Difficulties Questionnaire; d Note, a detailed description of this measure is provided in the next chapter; e PHQ-9 = Patient Health Questionnaire-9; f GAD-7 = Generalised Anxiety Disorder-7

2.3.9 Randomisation

Participants were randomly allocated (1:1) to either the intervention (VIPP-SD; Juffer et al., 2008a, 2008b) or control (usual care) trial arm. The randomisation list was generated by an independent statistician and used block sizes of 2, 4, and 6 varying at random. Randomisation was stratified by recruitment site and number of participating caregivers (one versus two). Participating families were allocated to the next available treatment code. Access to the allocation sequence was restricted and not accessible to any members of the research team.

2.3.10 Intervention arm: VIPP-SD

Families randomised to receive the VIPP-SD programme were offered six home-based sessions of one to two hours duration, delivered fortnightly by a trained health professional. The intervention was delivered by 40 trained therapists. Each therapist received four days of training, undertook a clinically supervised practice case, and received ongoing clinical supervision. Most therapists were health visitors (25%), nursery nurses (20%), and clinical psychologists (17%) with smaller numbers from research or psychiatry backgrounds. Visits were split into two parts. In the first part, the caregiver and child were filmed during 'everyday' interactions (e.g., reading a book, playing with toys, a mealtime). The therapist then used the video recordings of these interactions to write a script to be used to guide feedback delivered at the following session. Feedback centred on the sensitivity and discipline themes prescribed in the VIPP-SD manual for that session (Juffer et al., 2008a). The feedback was 'personalised' in that it was unique to moments shared between caregivers and their children. The second half of the home visit was spent reviewing the clips taken in the previous visit, with the therapist pausing the video every 10 to 30 seconds to deliver feedback from the pre-prepared script. Fidelity was assessed by rating a random selection of audio-recorded visits for adherence to the manual. Of those selected, 94% met the threshold for adherence. In families where two caregivers were participating, those randomised to the treatment arm (n = 26) were offered a modified version of VIPP-SD for coparents. VIPP-Co (Video-feedback Intervention to promote Positive Parenting for Co-Parents; Iles, Rosan, Wilkinson, & Ramchandani, 2017) follows the same model as VIPP-SD but includes an extra focus on the triad (both caregivers and child) interacting together.

2.3.11 Control arm: Usual care

Participants in both the VIPP-SD and usual care arms continued to access treatment as usual. In most cases, this care is minimal as there are no standard NHS care pathways for early onset behaviour problems. General support available to families was mostly through health visiting and GP services. Data on health and social care service use was collected through an interview with caregivers at each time point (CA-SUS; Byford et al., 2015).

2.3.12 Blinding

Research assistants were 'blind' throughout data collection and analysis. In the rare instances where unblinding did occur (2% and 4% of assessments at each follow-up point), recordings of researcher-rated measures were double scored by a second, blinded assessor.

2.3.13 Participant retention

The recruitment and retention of young families over a period of months and years poses a significant challenge for research teams. It is estimated that less than 31% of publicly-funded UK trials reach their original recruitment target (McDonald et al., 2006) with average attrition rates of 20 to 32% in paediatric intervention studies (Karlson & Rapoff, 2009). To maximise the reliability and validity of study findings, multi-strategy approaches to recruitment and retention are necessary. Schoeppe, Oliver, Badland, Burke, and Duncan (2014) emphasise the need to: design feasible studies that engage stakeholders throughout planning and implementation; build trustful relationships between researchers and families; implement flexible data collection procedures that are convenient for families and minimise burden; undertake extensive and persistent follow-up procedures; use strong leadership to develop cohesive research teams; and use researchers who have a sense of humour and an ability to deal with issues arising during data collection and intervention implementation.

HSHS was conducted in partnership with the trial's Patient and Public Involvement group. All aspects of data collection were carefully piloted to ensure the process was not overly burdensome for families. We had a flexible approach to data collection, often working evenings and weekends, and, where they were unable to complete a full home visit, families were offered the opportunity to complete study measures over the telephone or online instead. To improve participant engagement, we created a HSHS website, families were sent study updates through newsletters, children were given certificates, sticker progress trackers, and birthday cards, and we shared short video clips with families as a keepsake. We achieved very high levels of retention in the trial with 95% of families completing their 5-month follow-up

assessment (n = 286) and 94% completing their 24-month follow-up assessment (n = 282). In total, some form of follow-up data was available for 98% of participating families.

2.3.14 Data monitoring and quality assurance

2.3.14.1 Data monitoring and cleaning

Data was monitored in accordance with the standardised operating procedures set out by the Imperial Clinical Trials Unit. All participant consent forms, 20% of key HSHS outcomes (PPACS, CBCL, and SDQ), and 10% of all other outcomes were source verified.

2.3.14.2 Missing data

For the PPACS, missing answers to individual questions were imputed based on gender and age of child, intervention arm, and information gathered from items in the rest of the interview. For the other measures (CBCL, SDQ, PHQ-9, GAD-7, and the Parenting Scale) missing items were dealt with by scaling up. The observed total score for each scale was divided by the scale's inflation factor (the proportion of the scale that had been completed). Overall, missing data throughout the trial was negligible.

2.3.14.3 Approach to statistical significance

Under the null hypothesis significance testing paradigm there has been a keen focus on determining whether research findings are 'statistically significant' (typically defined as p-values less than the threshold of 0.05). However, there have been calls to demote statistical significance to a position where it is considered alongside other important statistical and non-statistical factors (e.g., Halsey, 2019). A p-value alone fails to provide an idea about the size of an effect or the importance of the result. Wherever possible, the study findings in this thesis are presented with an appropriate effect size and the magnitude of effect is commented on.

2.3.15 The use of Healthy Start, Happy Start data in this thesis

The HSHS study was principally designed to address the primary hypothesis that a video-feedback parenting intervention had the potential to reduce behaviour problems in one- and two-year-olds. However, this chapter serves to demonstrate the rich data source the trial provides for the exploration of other important and intriguing research questions. In the next three chapters, I utilise the longitudinal and experimental design of this study to better understand the behavioural development of very young children, the influences of parenting and a parenting intervention on family functioning, and how children themselves can enhance our understanding of these developmental pathways.

CHAPTER 3: Exploring the use of a child-informant measure of child behaviour in a sample of three- and four-year-olds

3.1 Chapter rationale

This chapter considers ways in which very young children can be involved in the reporting of their own behaviour and experiences in research using play- and narrative-based measures. Although the value of such measures is being increasingly recognised, their use has rarely been extended down to children under five years of age, particularly in samples considered to be more clinically vulnerable. This study explores the utility of a brief dolls play story stem procedure in a sample of three- and four-year-olds with elevated externalising behaviour difficulties. I consider the feasibility of delivering such a measure with this sample, explore associations between children's narrative responses and their gender, age, and parent-related variables, and compare the information collected through the story stem assessment to data obtained through parent- and teacher-reported measures of child behaviour.

3.2 Introduction

Children's capacity to bring rich and valuable perspectives to developmental research has been more widely recognised in recent years. We now know that, when given authentic opportunities and effective tools to do so, children can provide meaningful accounts of their experiences (O'Farrelly, 2021). However, the extension of such opportunities to preschool-aged children is less common and oftentimes met with uncertainty and scepticism. This is despite the fact the necessity to integrate children's own perspectives in matters which affect their lives starting in the earliest years was mandated by the *United Nations Convention on the Rights of the Child* over three decades ago (United Nations Committee on the Rights of the Child, 2005; United Nations Convention on the Rights of the Child, 1989). More recently, the European Commission (2021) stipulated this requirement in the EU Strategy on the Rights of the Child, a framework developed in consultation with 10,000 children. Despite these advances in children's rights, a widespread absence of their perspectives in research remains.

In this section, I reflect upon the potential of play-based instruments in accessing and understanding young children's experiences. Then, I outline a set of considerations that merit careful thought when undertaking research with children. Finally, I discuss the use of story stem tasks that utilise doll play methodologies, before presenting the study's aims.

3.2.1 Children as informants

3.2.1.1 Assessing children's behaviour and experiences

In Chapter 1, I highlighted the valuable opportunity early intervention offers in mitigating some of the far-reaching effects of behaviour problems. However, a key component of successful intervention is the early and accurate identification and understanding of children's difficulties. Studies of child behaviour are largely restricted to the unilateral view provided by primary caregivers (Luby et al., 2007). These reports can be confounded by reporting biases including limitations to parents' understanding of normative child functioning and their own experiences of psychopathology (Ringoot et al., 2015). Single informants may also be biased towards behaviours they regard as significant, or towards those they expect (e.g., gender-specific cultural expectations), leading to an overemphasis or fixation on these behaviours while overlooking others (Najman et al., 2001). Second informant data (e.g., second caregivers, teachers) can partially overcome some of these issues but is often limited by low response rates in research contexts (Arseneault et al., 2005).

Typically, a low to moderate degree of convergence is seen across multi-informant data measuring child behaviour, with a meta-analysis of 119 studies finding that agreement between caregivers, educators, clinicians, and trained observers averaged $r = .28$ (Achenbach et al., 1987). De Los Reyes and Kazdin (2005) noted that "informant discrepancies have been found in virtually every method of clinical assessment" (p. 483) used to measure child behaviour. One interpretation of this low-to-moderate degree of agreement is that informants each provide valuable but different information about the child's functioning across varied contexts, reflecting the situational-specificity of the informant-child relationship (Achenbach, 2006; Arseneault et al., 2005). Discrepancies are also thought to be influenced by the problem being reported on, whether the behaviours are directly observed or inferred, whether or not they occur across multiple settings, and whether the exhibited behaviours are mild or severe (Achenbach, 2006; Hinshaw, Han, Erhardt, & Huber, 1992; Kolko & Kazdin, 1993). Varying familial and contextual factors can also influence reports. For instance, Offord et al.'s (1996) study of parent- and teacher-reported child behaviour revealed that family functioning and parental psychopathology were associated with caregiver reports, while child gender and family socioeconomic status were associated with teacher reports. Multi-informant data can be particularly useful when measuring constructs for which there is no single, objective 'gold standard' measure. Arseneault et al. (2005) emphasised that "all informants' reports are imperfect measures of children's behaviour…collecting information from multiple sources is necessary for a comprehensive assessment of children's behaviour problems" (p. 547).

Including children themselves in information gathering may offer a unique and unrestricted perspective of their behaviour across settings (Arseneault et al., 2005; Kraemer et al., 2003). The intention of such an approach is generally not to supersede the perspectives offered by adults but to offer a complement to their reporting. However, these efforts must come equipped with developmentally appropriate tools and practices. Ways to facilitate children's involvement in developmentally sensitive ways are discussed next.

3.2.1.2 Facilitating children's involvement in the assessment of their development

Researchers intending to facilitate children's involvement in research must recognise the ways in which children's motivations, language, cognitions, and behaviour differ to those of adult participants. A review from Deighton et al. (2014) identified 11 child-reported measures of mental health and wellbeing in samples of 11- to 18-year-olds, including child-reported versions of the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001), Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), and the Behaviour

Assessment System for Children (BASC; Reynolds & Kamphaus, 1998). Modifications to these measures focused on the phrasing of items. This can help account for children's emerging language skills but does little to accommodate for the attention span and cognitive abilities of younger children (Arseneault et al., 2005).

Indirect or representational approaches commonly use play and storytelling to invite children to project their experiences through engaging stimuli (e.g., pictures, puppets, and dolls). Play-based methodologies have been used to enhance our understanding of varied aspects of children's development. These include the use of pictures in interviews of mental health symptomology (Valla, Bergeron, & Smolla, 2000; 6-11 years) and school adjustment (Murray & Harrison, 2005; 4-6 years), as well as puppet interview approaches as self-reported measures of behavioural difficulties (Measelle, Ablow, Cowan, & Cowan, 1998; 4-8 years) and children's playfulness (Fink, Mareva, & Gibson, 2020; 5-7 years).

Play and storytelling are familiar to children and correspond with natural ways that they engage with their worlds. Their use in research accommodates for children's development in several ways. First, the engaging nature of play may improve children's capacity to attend to the measure (Goldstein, 2012). Second, children may feel able to express themselves more freely than when under direct questioning, particularly as children often spontaneously represent salient aspects of their experiences in play (Harrison & Murray, 2015; Murray et al., 1999). Third, play is a form of communication even very young children tend to be well-versed in. Language development occurs at an extraordinarily rapid rate in the preschool years and there is considerable variation in children's verbal proficiency (Beltz, Blakemore, & Berenbaum, 2013). Using play may help to break down limits put on children's ability to participate when participation is contingent on verbal responses alone. However, there are relatively few valid and reliable play-based approaches that facilitate preschool-aged children's involvement in research processes.

3.2.2 Doll-play story stem narrative assessments

Narrative-based dolls play assessments involve encouraging children to manipulate a set of figures while telling a story (Buchsbaum & Emde, 1990). These techniques are said to offer a developmental bridge between behavioural observations, typically used with infants, and strictly verbal interviews (Bettmann & Lundahl, 2007). Most children acquire the skills for narrative production by three years old, meaning such methods are appropriate for use with preschool-aged populations (Bretherton & Beeghly, 1982; Buchsbaum & Emde, 1990).

The method's early roots lie in psychoanalytic therapy but doll play as a systematic assessment procedure has branched out across research and clinical contexts over the last several decades (Yuval-Adler & Oppenheim, 2015).

Story stem completion tasks involve an adult interviewer delivering an emotionally evocative story beginning – commonly a dilemma – before inviting the child to complete the story using words and play-action. The MacArthur Story Stem Battery (Bretherton & Oppenheim, 2003; Bretherton, Oppenheim, Buchsbaum, & Emde, 1990) and the Attachment Story Completion Task (Bretherton, Prentiss, & Ridgeway, 1990) were two of the earliest systematic story stem assessment tools. The developers encouraged others to adapt and develop new stems and coding frameworks. This reduces comparability between studies but broadens the scope of research interests these methods cater for. Such adaptations include the Manchester Child Attachment Story Task (Green, Stanley, Smith, & Goldwyn, 2000), which focuses on classifying attachment relationships of school-aged children, and the Doll's House Play approach (Murray et al., 1999), which prescribes a less structured format to elucidate aspects of family experiences in an open way. Providing a highly detailed discussion of the use and properties of these measures is beyond the focus of this study. Instead, Table 7 outlines key information related to several measures, informed by a review conducted by Bettmann and Lundahl (2007). The training to administer and code these measures can be resource intensive, but the batteries are flexible, portable, and relatively simple to administer. The empirical evidence base supporting the use of such tools reports good inter-rater reliability and validity (see Bettmann & Lundahl, 2007; Page, 2001; Yuval-Adler & Oppenheim, 2015 for reviews).

In these methods, dolls, teddies, animal figurines, and Duplo or Playmobil figures have been used as a stimulus (Bretherton, Oppenheim, et al., 1990). Getz, Goldman, and Corsini (1984) found that three- to five-year-olds were better able to engage with hypothetical interpersonal problems and offered more extensive responses when using three-dimensional props, compared to two-dimensional pictorial stimuli or words alone. This effect was particularly strong for impulsive children. One suggestion offered for this finding was that young children are better practiced at manipulating objects than words. This may explain why children tended to initially favour 'acting out' their responses before 'translating' them into verbal form in Getz et al.'s (1984) study. The authors propose that props more effectively communicate with the representational, non-verbal memory storage of young children, better facilitating the child's understanding of the situation and their thought and language

production. Extending upon this, Mize and Ladd (1988) found that enactment explained significantly more variance in pre-schoolers' prosocial and aggressive behaviours than pictorial stimuli.

In sum, play- and story-based methods may offer a particularly engaging way to gain the perspectives of very young children. Next, I consider the kinds of understanding that can be drawn from children's responses to such methods.

Table 7

Overview of Measures Developed to Explore Children's Narrative Representations

Measure	Procedure	Key themes	Number	Key studies	Age
Attachment Doll-Play Interview	Interviewer delivers stems focused on separation and reunion between children and their mother.	Emotional openness, resolution, and emotional tone	6 stems	Oppenheim (1997)	3-5yrs
Attachment Story Completion Task	Interviewer uses dolls to deliver five stems related to family-focused scenarios.	Security of attachment	5 stems	Bretherton, Ridgeway, et al. (1990)	3yrs+
Dolls' House Play	The child is encouraged to use dolls to act out a mealtime, bedtime, a bad and nasty time, and the best and favourite time.	Parental care, neglect, and hostility, caregiving, and narrative structure	4 stems	Murray et al. (1999)	5yrs
MacArthur Story Stem Battery	Interviewer uses dolls to deliver an emotive story stem before asking the child to "Show me and tell me what happens next".	Family relations, parent-child attachment, child behaviour, social, emotional, and moral development	14 stems	Bretherton, Oppenheim, et al. (1990); Bretherton and Oppenheim (2003)	3yrs+
Manchester Child Attachment Story Task	Interviewer delivers story beginning, child is asked to use dolls in a dolls' house to act out story endings.	Attachment-related behaviours, disorganised or disorientated attachment, narrative coherence	6 stems	Green et al. (2000)	5-7yrs
Symbolic Play Assessment	The interviewer delivers scenarios, mostly related to "illness" and the child is encouraged to play through the scenarios.	Parental anxiety, parental response, child anxiety, serious outcomes, and use of language	5 stems	Ramchandani et al. (2011)	6-9yrs
Teddy Bears' Picnic	Researcher presents a family of bears. Child is assigned to brother or sister bear and asked to respond to stems.	Family relationships, portrayal of helping, caring and nurturing and hostile, selfish, and rejecting	6 stems	Mueller and Tingley (1990)	4-6yrs

3.2.2.1 What can narrative representational measures tell us?

Story stem completions are reliant on children's capacity to coherently organise their thoughts, regulate their emotions, and draw on their own experiences and ideas. It has been suggested that responses not only tell us about children's behavioural tendencies but also offer a 'window' into their inner experiences and representations of themselves and others (Bettmann & Lundahl, 2007). There is significant variance in the ways children's narratives have been coded. This is unsurprising given that "the story stem approach does not represent a single measure designed to assess a single construct" (Yuval-Adler & Oppenheim, 2015, p.367) but rather a tool to elicit children's narratives for a range of purposes. However, three common dimensions of analysis are seen across the literature (Page, 2001; Yuval-Adler & Oppenheim, 2015):

1. The organisation and structure of storytelling (e.g., narrative coherence, avoidance).
2. The content of stories (e.g., themes related to behaviour, attachment, caregiving).
3. The child's behaviour during the task and their interactions with the interviewer (e.g., responsiveness to the interviewer, involvement with the measure).

Story stems are unlikely to elicit literal replays of children's experiences. However, it has been suggested that, while children's play does not reflect an exact copy of reality, it provides insight into "how the *child* reflects *upon* reality" (Hodges & Steele, 2000 p. 435). The research that documents associations between children's narratives and their observed or adult-reported behaviour provides confidence in the idea that children's stories are not 'pure fantasy' and that children have the capability to act as valid informants of their own behaviour (Bretherton & Munholland, 2008). A portion of these studies are now considered.

Children's attachment representations are one of the more well-studied constructs in the story stem field (Yuval-Adler & Oppenheim, 2015). Many of these studies have focused on the ways in which children represent key attachment figures in their stories and the coherence with which children can tell their narratives – a construct that is considered to be linked to the organisation of children's attachment representations (Sher-Censor & Oppenheim, 2004; Steele & Steele, 2005). For instance, in low-risk samples of five- to eight-year-olds, Moss, Bureau, François, Béliveau, Zdebik, and Lépine (2009; $N = 104$), Dubois-Comtois, Cyr, and Moss (2011; $N = 83$) and Gloger-Tippelt, Gomille, Koenig, and Vetter (2002; $N = 28$) all established links between children's attachment-related behaviours and their narrative coherence. Sher-Censor and Oppenheim (2004; $N = 113$) also reported that in a low-risk

group of four-and-a-half-year-olds, coherence of narratives and, in a subset of stories, aggressive and destructive themes were related to attachment style assessed three and a half years earlier. However, our understanding from these studies only extends to low-risk, school-aged samples. Further to this, attachment security represents only one important component of the parent-child relationship. We know less about associations between other aspects of parenting (e.g., discipline style, sensitivity) and children's narrative representations, especially with younger, more clinically vulnerable children.

Relatively few studies have focused on the thematic content of stories (e.g., aggressive, prosocial, or conflict themes). Of particular interest to the current study is research exploring associations between children's narratives and adult-reported behaviour problems (e.g., Warren et al., 1996, $N = 51$, 3-5 years, $r = .64$; Zahn-Waxler et al., 1994, $N = 89$, 4-5 years, rs .22-.25; Moss et al., 2009, $N = 109$, 8 years, rs .20-.27; von Klitzing & Bürgin, 2005, $N = 38$, 4 years, rs .30-.58; Page et al., 2011, $N = 46$, 6 years, Cohen's f^2 .07-.16; Woolgar, Steele, Steele, Yabsley, & Fonagy, 2001, $N = 74$, 5 years, rs .22-.26; von Klitzing, Kelsay, Emde, Robinson, & Schmitz, 2000, $N = 652$, 5-7 years, rs .09-.23). There is variance in the strength of association across these studies. von Klitzing and Bürgin (2005) and Warren et al. (1996) reported moderate to strong associations in small samples, while associations between child- and adult-reported data were of a lower magnitude in the other studies. As discussed earlier, this is not uncommon in the field of developmental psychology, with agreement not commonly exceeding $r = .20$ in child-adult informant studies (Achenbach et al., 1987; McConaughy, Stanger, & Achenbach, 1992). Broadly speaking, these findings indicate that there is some level of convergence between behaviours that adults observe and the representations children convey in response to story stems. However, these findings are almost exclusively derived from school-aged, non-clinical, low-risk samples of children, predominantly from middle-class, two-parent families. Most studies are also limited by small sample sizes, possibly due to the time-intensive nature of such data collection. There is a need to better understand the use of these measures in less well-studied populations of children.

3.2.2.2 Using story stems in under-represented groups

While much of the story stem research has been conducted with homogenous samples of children, there is a small body of literature that extends to children with a wider range of life experiences. For instance, small numbers of studies have included children who have experienced, or are at risk of, maltreatment and abuse (Buchsbaum, Toth, Clyman, Cicchetti, & Emde, 1992; Hodges, Steele, Hillman, Henderson, & Kaniuk, 2003; Toth, Cicchetti,

Macfie, & Emde, 1997), have been adopted (Hillman et al., 2020), are living in post-divorce families (Page & Bretherton, 2001), or have mothers in prison (Poehlmann, 2005).

Of particular interest to the current study is research exploring the use of narrative tools with children experiencing, or at an increased likelihood of, behavioural difficulties. Clinically-referred samples have been found to include more aggressive story content (e.g., Hill, Fonagy, Lancaster, & Broyden, 2007; 5-8 years; total $N = 66$, clinical $n = 41$) and their narrative representations have been found to correspond with parent-reports of behaviour problems (e.g. Green, Stanley, & Peters, 2007; 4-9 years; $N = 61$). Interestingly, von Klitzing, Stadelmann, and Perren (2007; 5 years; total $N = 187$, clinical $n = 80$) reported stronger associations between adult-reported child behaviour and children's prosocial story themes ($r = .38$) and narrative coherence ($r = .39$) in their clinically-referred subsample compared to their non-clinical subsample (rs .11-.18). These findings highlight possible differences in what these measures can tell us for clinical and non-clinical populations. The inclusion of children with more varied individual- and family-functioning is encouraging but remains largely restricted to school-aged children. Further to this, children exhibiting clinically significant difficulties often only formed a subsample of participants, with the remainder recruited from community samples. Finally, most adult reports of children's difficulties were limited to single reporters, mostly mothers, and mostly using checklist measures alone, providing a restricted view of children's development.

These issues restrict our understanding of the suitability of these methods for very young, clinically vulnerable children. Yuval-Adler and Oppenheim (2015) speculated that such methods may be less reliable in younger pre-schoolers, particularly in high-risk groups of children. The specific research this speculation is based on is a little unclear. It could be informed by findings included in their review from Warren et al. (1996) who reported that children's narrative coherence was not associated with behaviour problems in three-year-olds but was in four- and five-year-olds, and that narrative coherence was only stable over 12 months in the older groups of children. Our understanding of other developmental and demographic factors that may be associated with the construction and content of very young children's narratives also needs enhancing. For instance, while differences in story stem constructs have been associated with child gender (e.g., Zahn-Waxler et al., 1994; Zahn-Waxler, Friedman, Cole, Mizuta, & Hiruma, 1996) and parental psychopathology (e.g., Hill, Murray, Leidecker, & Sharp, 2008; Pass, Arteche, Cooper, Creswell, & Murray, 2012) these differences have not been well-considered in younger children. More research is needed to

explore if and how these approaches can benefit our assessment and understanding of the behaviour of children as young as three years of age.

3.2.3 Conclusion

The value of including children as research informants is multi-layered. Including data from multiple informants in the assessment of difficult-to-measure psychological constructs, of which there is no one source of 'correct' information, is encouraged (Kraemer et al., 2003). The small to moderate degree of convergence between informants has been seen as reflective of each reporter's unique perception of the child's development (Duhig, Renk, Epstein, & Phares, 2000). Taking account of multiple perspectives from children themselves, their caregivers, and teachers may enhance the accuracy and precision of the information we obtain (Wolcott, Williford, & Hartz-Mandell, 2019). Importantly, previous research indicates that adding children as informants may help to address biases in parent reports of their children's behaviour (Ringoot et al., 2015). However, studies of the use of child-informant measures with very young children (aged under five years), particularly those in more clinically vulnerable populations, are scarce. Identifying ways to better understand children's internal worlds, something that is challenging to gain direct access to, in an effective and efficient manner may offer valuable opportunities for research and clinical care, to integrate children's perspectives into what are often adult-focused assessments of children's behaviour.

3.2.4 Present study research aims

The present study aims to contribute to addressing these gaps by adapting a story stem measure of child-reported behaviour and considering its utility in a sample of three- and four-year-olds recruited based on early behavioural difficulties. The study will focus on the feasibility of this measure, before considering if and how the information gathered relates to concurrently-recorded data from parents and teachers. Given the exploratory nature of this analysis, specific hypotheses are not made. However, the key research aims are as follows:

1. To explore the feasibility of delivering the story stem measure, by examining completion rates and reasons for incompletion.
2. To explore associations between adult-rated (parent and teacher) child behaviour and children's own reports using the story stem method.
3. To explore associations between children's gender, age, and narrative representations.
4. To explore associations between parental education and symptoms of depression and anxiety and children's narrative representations.

3.3 Method

3.3.1 Design

Data for this study were collected in the broader context of the Healthy Start, Happy Start trial (HSHS; Ramchandani et al., 2017). A detailed overview of the HSHS trial's design, recruitment, and sample is provided in Chapter 2.

The current research has a cross-sectional, observational design, using data collected at the final HSHS follow-up assessment when children were three to five years old. Its primary focus, the use of a story stem assessment battery, was included in the 24-month follow-up assessment as an exploratory outcome. I was involved in all aspects of the adaptation of the story stem measure and the data collection and coding of the data obtained.

3.3.2 Participants

Participants were 261 families (child and primary caregiver) who completed their 24-month follow-up assessment in the HSHS trial, when children were three to five years old. Children had been recruited into the trial based on parental reports of elevated externalising behaviour problems two years earlier at 12 to 36 months old. Further information about eligibility, screening and recruitment of participants is outlined in Chapter 2, Section 2.3.

3.3.3 Measures

All measures used in this study were collected at the 24-month follow-up time point except for the demographic interview which was administered during families' baseline assessment. The following section provides detailed information about the doll's house play story stem assessment battery. It also offers a broad overview of measures of adult-reported child behaviour, parental discipline practices, and psychopathology. More detailed information about these measures is provided in Chapter 2, Section 2.3.8.

3.3.3.1 Child informant measure: A doll's house play story stem battery

A narrative-based doll's play assessment was administered to elicit children's behavioural representations (prosocial and disruptive behaviours) and to provide a measure of the coherence with which children told their stories. The battery also allowed for an observational assessment of the child's behavioural dysregulation during the completion of the seated, instruction-based activity. This section describes the development and administration of the measure and details the approach undertaken in coding the data.

3.3.3.1.1 Development of the measure

The administration and coding of the story stem measure developed for use in HSHS was informed by two existing approaches that use doll play techniques: the MacArthur Story Stem Battery (MSSB; Bretherton & Oppenheim, 2003; Oppenheim, Emde, & Warren, 1997) and the Dolls House Play approach (DHP; Murray et al., 1999; Ramchandani et al., 2011; Woolgar & Murray, 2010). Murray et al.'s (1999) DHP method was reported to take 20 minutes to administer (excluding set-up and warm-up), while the MSSB was reported to take approximately 60 minutes (Emde, Wolf, & Oppenheim, 2003). Adaptations to these approaches were necessary for use in the current study as the measure was to be embedded within a broader research assessment that was two and a half hours in duration. To limit the burden on families, the aim was for this measure to be delivered in approximately 20 minutes, including time for set-up and warm-up.

Three story stems were chosen for use in the current study (see Table 8 for more detail). Two stems were adapted from the MSSB (Bretherton, Oppenheim, et al., 1990), and one from the Symbolic Play Assessment (Ramchandani et al., 2011; adapted from Murray et al., 1999's DHP). The stems were selected to elicit children's narratives and representations concerning dimensions of their behaviour, including behaviours related to conduct problems and prosociality. The MSSB developers encouraged researchers to develop new stems (following careful piloting) or to mix and match stems from other measures. They stressed that the MSSB was not developed as a standardised test but that researchers should follow the "battery's spirit or underlying aim of facilitating story telling" (Emde et al., 2003, p.66).

The approach in the current study of having the child pick a doll to represent themselves, their chosen family members, and best friend was in-line with the method undertaken by the Dolls House Play research group (Murray et al., 1999; Ramchandani et al., 2011; Woolgar & Murray, 2010). The aim of this particular approach was to allow for the child to become emotionally engaged in the story being told, potentially "revealing those aspects of family functioning that are most salient to him or her" (Ramchandani et al., 2011, p. 411). Some research has indicated that asking children to represent their own family may help them to identify with the dolls, thus leading to the production of more identificatory themes and/or the representation of autobiographical information (Woolgar, 1999).This contrasts with the approach prescribed in the MSSB which used pre-specified dolls and a 'standard' family (e.g. Susan, Mum and Dad, sister Jane, and friend Laura).

3.3.3.1.2 Piloting and training on measure administration

I led on the piloting of the story stem measure with 20 three- and four-year-olds. The first phase of piloting involved 15 children from two local nursery schools. In this phase, we found that the home props (refrigerator, cupboards, and sink) and park props (swing and slide set) appeared to distract from the child's storytelling (e.g., a preoccupation with getting the dolls to balance on the swings rather than engaging with the story prompt). Therefore, before phase two of piloting, these props were removed, leaving just a table, chairs, and a jug to represent the home and a green piece of felt and small foam ball to represent the park.

The second phase of piloting took place in the homes of five families as part of a wider piloting of the complete two-year follow-up assessment. This allowed us to further adapt the measure for use in the home context. In their own homes, children were more prone to bring in their own toys to the stories than we had seen in the nurseries. Guidance to minimise distractions within reach of the child, as well as scripted explanations that encouraged the child to only use the researcher's props, were added to the administration manual. The HSHS trial manager and principal investigator also reviewed videotaped recordings of these piloting sessions to offer feedback and advice.

The HSHS trial manager observed the three researchers carrying out data collection during two or three cases before we started administering the measure independently. Supervision sessions with the trial manager and principal investigator were conducted for the first seven months of task administration to ensure consistency in approach. Following these sessions, where videos of completed story stem assessments were used to guide discussion, small alterations to the measure's procedures were made (see Appendix E). Changes centred on improving guidance for researchers to optimise measure delivery. These included suggestions to ask the child's caregiver where the child most typically sat during seated activities or to place a cushion on the floor to 'root' the child if a table was unavailable. Another suggestion was for researchers to shift their approach to 'wonder aloud' for children who seemed to lack confidence in responding to direct questioning (e.g., replacing "how does Jack feel about his spilt juice?" with "hmm, I wonder how Jack feels about his spilt juice.").

3.3.3.1.3 Procedure for delivering the HSHS story stem battery

The home visit was split into two parts. During the first part, measures were mostly administered to the child's primary caregiver with the child only being involved in the collection of video-recorded parent-child interactions. In the second part of the visit (~40

minutes) the researcher focused on administering measures with the child, while their caregiver was completing questionnaires. By this point, the researcher had usually been in the child's home for two hours and children were typically very open to working independently with the researcher. Researchers were trained to remain vigilant to verbal and non-verbal signals of fatigue, disinterest, or distress in the child.

When administering the story stems, the researcher generally sat with the child at a table and, where possible, a little away from the caregiver to limit instances where caregivers would provide direction or answers for the child. The researcher explained that they were going to tell the beginning of the three stories and the child could tell the ending. If the child was happy to take part, they were shown fifteen dolls and asked to pick the dolls they felt were most like them and their family members. The measure began with a warm-up story (a party) which provided an opportunity to convey to the child what was expected of them in the activity. The warm-up's four aims were for the child to: speak to the researcher, pick up the dolls and manipulate them, speak for the dolls, and say something that related to the scenario at hand. In the warm-up, researchers could model the enactment for the child and deliver prompts to engage the child in the play (e.g., "I think Mummy's going to have a big bite of a cake, she says, 'Yum, that's delicious!' What does Jack say about the cake?"). However, direction and leading questioning was not used in subsequent story stems.

Once the child had fulfilled the four aims of the warm-up, they were invited to play through the three scenarios. Throughout story completion the researcher used a series of scripted non-leading prompts and standard probes (see Table 8). The dolls and props allowed for verbal and non-verbal communication in the child's storytelling and the children's responses were recorded using a video camera on a tripod.

Table 8

Story Stem Scenarios and Prompts Presented to the Child in the Story Stem Measure

Stem	Prop set up	Summarised story beginning told by researcher	Additional prompts used by researcher (some optional)
Warm up[a]		Before we tell our three stories, I think the family could have a party. I wonder what kind of party they could have.	• Modelling of narration and manipulation of dolls. • Demonstrations and leading prompts could be used here but not in subsequent story stems.
1. Spilt juice[b]		The family are thirsty and they're going to have some juice. Now [child's name] reaches across the table and uh-oh! He spills the jug of juice all over the floor! Show me and tell me what happens now.	• What happens about [child] spilling the juice? • Did anything else happen? • How does [child/primary caregiver] feel? Why? • Show me what happens in the end, how does it all end up?
2. Hurt knee[c]		[Caregiver(s)] are over here and [child] and [friend] are playing here in the park. Suddenly [child] runs and trips and falls and cuts their knee. "Ow! I've really hurt my knee!" Show me and tell me what happens now.	• What happens about [child's] hurt knee? • [Child] says "Ow, my knee is really sore." What happens about their hurt knee? • How does [child/primary caregiver] feel? Why? • Show me what happens in the end, how does it all end up?
3. Three's a crowd[b]		[Study child] is playing with Harry and Harry's new ball. Sam comes along and asks, "Can I play with you?" [Study child] says "Sure!" but Harry says "No way! I don't want to play with Sam." Show me and tell me what happens now?	• What happens about Sam wanting to play with the ball? • "But I said I don't want Sam to play with us." What happens about who gets to play? • How does [study child/excluded friend] feel? Why? • Show me what happens in the end, how does it all end up?
Wind down[a]		You can play with any dolls you want now.	• Child was encouraged to have fun using any props or dolls.

a Data not included in analysis;
b Adapted from MacArthur Story Stem Battery (Bretherton et al., 1990);
c Adapted from Doll's House Play (Murray et al., 1999)

3.3.3.1.4 Developing a story stem coding framework

The decision to complete video coding of the data, rather than analysing transcriptions of children's responses alone, was driven by the fact that both verbal and non-verbal communication and behaviour during play was of interest. The coding scales developed for the HSHS story stems were informed by coding approaches undertaken in other studies that have used children as informants. Professors Lynne Murray and David Oppenheim, leading experts in the doll's play field, were also consulted to improve aspects of the coding scheme.

The scoring was divided into three core sections: representations of child behaviour, the coherence with which the children told their stories (narrative coherence), and a rating of children's behavioural dysregulation during measure completion (see full coding scheme in Appendix F).

Behaviour-related themes. Children's stories were coded for aggressive, destructive, ostracising, and prosocial behaviours and themes. These four behaviours were selected for two key reasons. Firstly, as the children were recruited based on early signs of behavioural difficulties it was of interest to explore themes related to conduct problems in their storytelling two years later. Secondly, three well-validated measures of parent-reported child behaviour and one measure of teacher-reported child behaviour were also available, allowing for the exploration of these behaviours in a multi-informant way.

Under the MacArthur Narrative Coding system (Robinson, Mantz-Simmons, & Macfie, 1992), prosocial and conduct-related behaviours were scored on a presence/absence scale. This approach does not account for variance in the intensity of the exhibited behaviour. Given the screening criteria of the HSHS sample, it felt important to capture detail and variance in themes related to these behaviours. Following the piloting of different coding approaches, it was decided that behaviours would be scored on a scale of 0 to 3 for intensity in each story stem. Each score was accompanied by example descriptions of story content (e.g., for aggressive behaviours: 1 = pushing, rough play, verbal threats; 2 = kicking, hitting; 3 = character attacks another, hits over and over, aggression resulting in injury or death). These definitions were predominantly informed by descriptions of aggression, exclusion, and empathy/helping in the MacArthur coding system. Other measures of these behaviours in young children were also used to further define and refine the thresholds for scoring in the coding scheme. This included descriptions of conduct problems used in a parent-reported measure for preschool-aged children (PPACS; Sonuga-Barke et al., 1994; Taylor et al., 1991).

Narrative coherence. Narrative coherence relates to a child's ability to tell their story in an organised and coherent way. Coding for narrative coherence was informed by a scale developed as part of the MacArthur Narrative Coding System (Robinson et al., 1992). This way of rating narrative coherence considers how the child's response addresses the dilemma presented in the story stem, the degree of embellishment, and the presence of incoherent shifts in the narrative. The narrative coherence scale developed by Robinson et al. (1992) was designed for use with four- to six-year olds and spanned 11 points. After piloting the application of this scale to the responses of three- and four-year-olds, the decision was made to collapse the definitions down into a 0 to 5, six-point scale (see Table 9 for example descriptions of these scores). It proved difficult to apply the full range of the original 0 to 10, 11-point scale to sample videos in a way that consistently demarcated high, moderate and low scores. This could possibly be because the stories in the current sample were told by younger children than the original scale was intended for and thus the stories may have been consistently shorter or less detailed than you might see from five- and six-year-olds. The six-point scale retained most of the original definitions for operationalising the 11-point scale, but it proved more reliable in consistently distinguishing between children's responses.

Table 9

Exemplar Descriptors for a Selection of Scores for Children's Narrative Coherence

Score	Description
1	**Not coherent.** There are fragmented shifts in story line. Child does not return to original story stem or address the conflict.
3	**Moderately coherent.** Child exhibits an understanding of the conflict but only offers a partial resolution, narratives may appear fragmented or include incoherent shifts. Child may shift the focus of the story away from the original dilemma (e.g. sibling spills juice repeatedly instead) or indirectly by offering an easier solution (e.g. child walks away without ball).
5	**Very coherent.** A very coherent, logical, sequential series of events that are related to the story stem. Child may add to the story but does not change the original story stem. An understanding of the conflict and a resolution to the conflict are presented. There are no incoherent shifts in the story and there is a lot of embellishment.

Note. These are exemplar scores taken from a 0 to 5 scale. Complete scale is available in Appendix F.

Observed behavioural dysregulation. The observational scoring for behavioural dysregulation formed a direct assessment of the ways in which children were able to modulate their behaviour during the structured story stem activity. To my knowledge, the story stem battery has not been used in this way before so there was not an existing coding scale of dysregulation for application to the measure. Instead, the framework for this scale was adapted from the Preschool Self-Regulation Assessment Profile – Assessor Report (PSRA-Assessor Report; Smith-Donald, Raver, Hayes, & Richardson, 2007) which has been described as providing a "broad brush snapshot of children's regulatory skill" (Raver et al., 2012, p. 261). This scale was originally developed for use following the delivery of a structured battery of seven different activities assessing both 'hot' and 'cool' self-regulation in behaviour, emotion, and attention (Denham, Warren-Khot, Bassett, Wyatt, & Perna, 2012; Smith-Donald et al., 2007).

Given that the story stem battery was not primarily designed to measure self-regulation, and only involved the administration of a single activity, it was not appropriate to apply the PSRA – Assessor Report in its entirety. Some of the original items assessed the child's ability to wait patiently for new activities to begin, to plan and consider new instructions and tasks, and to look closely to make thoughtful choices before answering 'hard' questions. These items were not possible to operationalise in the context of the story stem task. Instead, following piloting of the scale, specific items measuring attention (e.g., sustains concentration), impulse control (e.g., refrains from indiscriminately touching test materials), and activity level (e.g., remains in seat appropriately during task) were selected. Items were scored on a scale of 0 to 3 with behavioural descriptions acting as anchors in the guide for scoring. Descriptions were informed by the original report but were tweaked for operationalisation in the current activity (e.g., references to grabbing at or damaging "blocks or paper" were changed to "dolls, bags of dolls, and doll's house equipment").

3.3.3.1.5 Procedure for training coders

I trained three graduate students to support in the completion of data coding. During training, we watched children's story responses together and used a stop-start approach to consider how to apply each scale to the data. Discussions focused on discrepancies or differences in the understanding and interpretation of the scales. The guidance in the coding framework was expanded upon or simplified where necessary. As training progressed, coders scored the videos independently and then discussed ratings to standardise the scoring of behaviours. Videos used in training were randomly selected HSHS videos, as I did not have

access to any other example story stem videos. Data used as part of training was formally coded at the end of the coding phase to ensure as much time as possible had passed since viewing in training. Length of training varied between 16 and 32 hours, as one coder was trained on a single construct while the other two coders were trained on three constructs. Each coder and I independently scored a set of 12-15 randomly selected videos to form a reliability set for each scale.

3.3.3.1.6 Establishing inter-rater reliability

The intra-class correlation coefficient (ICC), an index of measurement that reflects both the degree of correlation and agreement between coders (Koo & Li, 2016), was used to measure inter-rater reliability. ICC estimates and their 95% confidence intervals were calculated using SPSS version 25 based on a single-rater, absolute-agreement, two-way mixed effects model for scales of narrative coherence, behavioural representations and observed behavioural dysregulation. While there is no single value for 'acceptable' reliability, there are widely used thresholds. Values below 0.5 are considered to be indicative of poor reliability, values between 0.5 and 0.75 indicative of moderate reliability, values between 0.75 and 0.90 indicative of good reliability, and values greater than 0.90 indicative of excellent reliability (Portney, 2020).

In the current study, all but one ICC calculation surpassed the pre-specified threshold for reliability (0.70) with estimates ranging from 0.75 to 0.98 (see Appendix G for full ICC results). However, due to a particularly discrepant case, there was one estimate for narrative coherence between two coders of 0.68. In ideal circumstances, further training would have been completed to ensure all reliability scores were above 0.70. However, due to the practical implications of the COVID-19 pandemic, the decision was made to begin coding of the main dataset in the limited time students had to complete analysis, particularly given that the value was very close to 0.70.

3.3.3.1.7 Allocation of videos for coding

To minimise any potential effects of bias, it was decided that I would not complete coding of any of the story stem assessments I had conducted (n = 95; 36%). It should be noted though that over the course of the HSHS trial I met 193 (64%) participating families through at least one assessment visit. Given this, I cannot be considered an entirely independent coder, although I was blind to group allocation. The three other coders were not involved in data

collection and were blind to all participant characteristics. Graduate coders coded between 59 and 148 cases each and I coded the remaining 108 cases (42%).

3.3.3.2 Demographic measure

Participants' sociodemographic information was collected using a researcher-led interview with the child's primary caregiver. Demographic variables included caregiver age, sex, ethnicity, educational attainment, employment status, and relationship status, and child age, sex, and ethnicity.

3.3.3.3 Parent-reported child behaviour measures

Given the exploratory nature of the current study, child behaviour was assessed using three different measures: a semi-structured, clinical interview and two caregiver-reported questionnaires. The psychometric properties of these measures have already been detailed in Chapter 2, Section 2.3.8 but a broad overview is also offered here.

3.3.3.3.1 Preschool Parental Account of Childhood Symptoms

Children's externalising behaviours were assessed using a modified version of the Preschool Parental Account of Childhood Symptoms (PPACS; Sonuga-Barke et al., 1994; Taylor et al., 1991). The PPACS is a semi-structured, researcher-led clinical interview that has been utilised in a number of epidemiological and intervention studies and found to have good psychometric properties (Sonuga-Barke et al., 1994; Taylor et al., 1986b; Taylor et al., 1986a). For use in the current study, the PPACS was modified so that it could be applied to the behaviour of infants and toddlers. Questions related to behaviours not often seen in very young children (including lying, stealing, and being cheeky or answering back) were removed and exemplar behaviours used to guide scoring were revised to reflect developmental expectations in younger children. The reference timeframe of interest was also altered from the last six months to the last four months to accommodate for the rapid rate of behavioural development in very young children.

The interview questions focused on behaviours related to ADHD/hyperkinesis (i.e. inattention, hyperactivity, and fidgeting) and conduct difficulties (i.e. temper tantrums, defiance/disobedience, destructiveness, and aggressiveness). The child's primary caregiver was asked to give detailed examples of the child's behaviour in different situations within the last week. These examples acted as anchors to the parents' memory and recall, allowing the interviewer to use their clinical judgement to rate the described behaviour according to

operationalised criteria. Caregivers were then asked to report on how representative the described behaviour was of the child, both in terms of severity and frequency, over the last four months, to ensure examples were typical and characteristic of the child over a longer timeframe. The interview generally took 45 to 60 minutes to complete. Interviewers undertook two days of face-to-face training in the measure, completed two or three supervised cases before conducting interviews independently, and received supervision on conducting and scoring the interview throughout data collection. Interviews were audio-recorded and 10% of interviews were randomly selected to measure inter-rater reliability. One-way, single-measurement, absolute agreement intra-class correlation coefficients indicated moderate to excellent levels of inter-rater reliability between assessors (ADHD subscale .98, conduct subscale .72, total score .92) at the 24-month follow-up.

3.3.3.3.2 Child Behaviour Checklist

The Child Behaviour Checklist (CBCL/1½-5; Achenbach & Rescorla, 2000, 2001) measures a broad range of behavioural, emotional, and social problems in young children. Respondents are asked to rate 100 items based on how true the described behaviour is of their child over the last two months on a three-point scale (0 = not true, 1 = somewhat or sometimes true, 2 = very true or often true). Scoring provides a total difficulties score, as well as an externalising and internalising score. This is a widely used and well-validated measure for the assessment of child behaviour problems and is suitable for use from toddlerhood. The test-retest reliability, concurrent validity, and internal consistency of the measure have all been reported at an acceptable level (Achenbach & Rescorla, 2001; Ivanova et al., 2010; Rescorla et al., 2011).

3.3.3.3.3 Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is widely used in research, clinical, and educational practice (Goodman & Scott, 1999). The SDQ includes 25 items designed to capture problems related to five subscales: conduct problems, emotional symptoms, hyperactivity/inattention, peer relationship problems, and prosocial behaviour. Respondents are asked to rate each item on a three-point scale (0 = not true, 1 = somewhat true, or 2 = certainly true) based on the child's behaviour during the last six months. In the current study, a version of the SDQ adapted by the measure's developers for use with younger children was used. Adaptations included revising the wording of three items to improve their developmental suitability (focused on lying, stealing, and impulse control). The psychometric

properties of the revised pre-schooler measure have been explored less than the 4-17 year old version, but the available studies have reported good psychometric properties related to internal reliability, factor structure, predictive validity, and convergent validity (Croft et al., 2015; D'Souza et al., 2017; Ezpeleta et al., 2013; Klein et al., 2013)

3.3.3.4 Parenting practices

3.3.3.4.1 Parenting Scale

The Parenting Scale (Arnold et al., 1993) measures parental discipline styles. The 30-item, self-report questionnaire asks respondents to report on their responses to child misbehaviour over the last two months. Situations are presented using seven-point Likert scales with an 'effective' and 'ineffective' approach to addressing the situation at each end. A total score and three subscale scores for laxness (permissive discipline), overreactivity (harsh discipline), and verbosity (overly wordy responses to misbehaviour) are generated. The Parenting Scale has been reported as having acceptable to good test-retest reliability, internal consistency, and concurrent validity (Arnold et al., 1993; Lorber et al., 2014; Prinzie et al., 2007). Since the measure's development, a two-factor structure of lax and overreactive parenting has been supported but studies have failed to replicate the verbosity factor found in the original development study (Harvey et al., 2001; Reitman et al., 2001; Rhoades & O'Leary, 2007; Steele et al., 2005). For this reason, the laxness and overreactive subscale scores were analysed separately in the current study and the verbosity scale was excluded.

3.3.3.5 Measures of parental psychopathology

3.3.3.5.1 Parent low mood

The Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) measures symptomatology and severity of depression. The nine-item questionnaire asks respondents how often they were bothered by nine symptoms over the last two weeks on a four-point scale from 0 (not at all) to 3 (nearly every day). The diagnostic validity, internal consistency, test-retest reliability, criterion validity, and concurrent validity of this measure has been well-established (Spitzer et al., 1999; Spitzer et al., 2000).

3.3.3.5.2 Parent anxiety

The Generalised Anxiety Disorder Questionnaire (GAD-7; Spitzer et al., 2006) measures symptomatology and severity of anxiety. The seven-item questionnaire asks respondents how often they have experienced the listed symptoms in the last two weeks,

ranging from 0 (not at all) to 3 (nearly every day). The internal consistency, test-retest reliability, and convergent validity of the measure has been found to be good to excellent (Spitzer et al., 2006).

3.3.4 Study procedure

The study assessments were conducted in participants' homes or, in a small number of cases, a private room in a community centre. Written informed consent was provided by participating caregivers for each phase of the study. The participant information sheet and the consent form (see Appendix D) were explained in person to participating caregivers. More detail around informed consent is provided in Chapter 2, section 2.3.6. The 24-month follow-up assessment took approximately two and a half hours. Following consent, the child's primary caregiver completed two interviews with the researcher (including the PPACS). Next, a set of parent-child, play-based interactions were video-recorded. Caregivers then completed a questionnaire battery (including the CBCL, SDQ, PHQ9, GAD-7 and Parenting Scale) while the researcher administered direct assessments of the child (including the story stem assessment).

3.3.5 Statistical analysis

The broad aim of this study was to provide insight into the use and utility of a brief story stem measure, assessing very young children's self-reported behaviours, as a potential supplement to adult-reported outcomes often collected in clinical and research settings. The study's analysis was exploratory in nature and aimed to:

1. Explore the feasibility of delivering the story stem measure.
2. Explore concurrent associations between adult-reported child behaviour and constructs derived from children's responses to the story stems.
3. Explore whether child behaviour, measured through the story stem assessment, varied based on parent-related (psychopathology, educational attainment, discipline practices) and child-related (age and gender) characteristics that typically predict (parent-reported) child behaviour.

3.3.5.1 Planned statistical analysis

First, the sample's sociodemographic characteristics were examined. Next, the descriptive statistics of key study variables were generated, and boxplots, histograms,

normality plots, and tests for skew and kurtosis were used to examine the distribution of study variables and detect outliers. Data completeness and reasons for missing data were then examined. Qualitative examples of children's responses to the story stem measure were produced for illustrative purposes. Following this, bivariate (Pearson) correlation coefficients were used to explore concurrent associations between parent- and teacher-reported behaviour and behavioural outcomes derived from the story stem measure. Next, concurrent associations between the story stem variables and factors commonly related to (parent-reported) child behaviour were explored, including associations with parental psychopathology, parental discipline, and parental educational attainment. Finally, any effects of child age and gender on story stem variables were run using *t*-tests and correlational analyses. Approach to missing data can be found in Chapter 2, section 2.3.14. Analysis was completed using SPSS version 25.

3.3.5.2 Sample size and power analysis

Estimating the statistical power that a study's sample size provides is important to understand the probability of making a Type II error. The G*Power tool (Faul, Erdfelder, Buchner, & Lang, 2009) was used to conduct a power analysis. This helped predict the necessary sample size to obtain sufficient power given an estimated effect size for each statistical test used. For correlational analysis, to detect a medium effect size ($r = 0.3$; Cohen, 1988, 1992), a sample size of 84 was required to detect an effect at 80% power ($\alpha = 0.05$, two-tailed). For an independent samples t-test, to detect a medium effect size ($d = .50$), a sample size of 128 was required to detect an effect at 80% power ($\alpha = 0.05$, two-tailed). The available sample surpassed these requirements.

3.4 Results

The structure of presented results is as follows: First, the recruitment and sociodemographic characteristics of the study sample are presented. Then, for Aim 1, I consider the feasibility of delivering the story stem measure with this sample of children. For this, completion rates and reasons for incompletion, as well as examples and frequencies of coded behaviours, are considered. For Aim 2, associations between the constructs derived from the story stem measure and parent- and teacher-reported child behaviour are explored using correlational analyses. For Aim 3, the relationships between the constructs obtained through the story stem measure and factors commonly related to child behaviour (when obtained through parental reports) are examined. These include caregivers' psychopathology, discipline practices, and educational attainment and child age and gender differences.

3.4.1 Description and exploration of study data

3.4.1.1 Recruitment and retention

Healthy Start, Happy Start (HSHS) participants included in the current analyses were 261 children who undertook the story stem battery, administered at the 24-month follow-up time point, and their caregivers. This sample represents 87% of the baseline sample of families randomised into the main trial two years earlier (N = 300) and 91% of the families who completed at least one outcome measure at the 24-month follow-up assessment (N = 286). Detailed information about the recruitment of this sample is provided in Chapter 2.

3.4.1.2 Sample characteristics

Children were between 36 and 65 months old (M = 48.0 months; SD = 7.1 months) at the time of the 24-month follow-up. There were slightly more male than female children participating (53% male), and most children were from a white (65%) or Mixed heritage (21%) background. Participating caregivers were predominantly female (96%), and all were either the child's biological or adoptive mother (96%) or father (4%). At baseline, most participating caregivers were married or living with a partner (87%). In total, 67% of the children's primary caregivers held a graduate-level qualification. See Table 10 for full demographic information. A detailed comparison of the Healthy Start, Happy Start study's sample and the general UK population was undertaken in Chapter 2, section 2.3.5. Tables 51 and 52 in Appendix H detail the sociodemographic profiles of families and descriptive statistics of study variables in this subsample compared to the full HSHS sample.

Table 10

Demographic Information for Participating Children and their Primary Caregivers

Characteristic	*n* (%), unless stated
Child characteristic	
Gender	
Female	122 (47)
Male	139 (53)
Age (months), mean (SD)	
Baseline	23.0 (6.7)
24-month follow-up	48.0 (7.1)
Ethnicity	
Asian	15 (6)
Black	14 (5)
Mixed	55 (21)
White	170 (65)
Other	7 (3)
Primary caregiver characteristic	
Gender	
Female	251 (96)
Male	10 (4)
Age (years), mean (SD)	34.7 (5.5)
Relationship to child	
Biological mother	250 (96)
Biological father	9 (3)
Adoptive mother	1 (0.5)
Adoptive father	1 (0.5)
Ethnicity	
Asian	27 (10)
Black	16 (6)
Mixed	18 (7)
White	191 (73)
Other	9 (4)

Characteristic	*n* (%), unless stated
Relationship status	
Married/civil partnership/cohabiting	227 (87)
Divorced/widowed/legally separated	5 (2)
Single and none of the above	21 (8)
In relationship but not cohabiting	8 (3)
Employment status	
Working for an employer	116 (44)
Looking after home and children	93 (36)
Self-employed	27 (11)
Paid parental leave	16 (6)
Full-time student	9 (3)
Highest qualification	
GCSE or lower	23 (9)
College (e.g., A levels, NVQ, BTEC)	62 (24)
Graduate	176 (67)

N = 261

3.4.1.3 Exploration of study variables

3.4.1.3.1 Descriptive statistics

Descriptive statistics of study variables, including means, standard deviations, median, range, inter-quartile range, and measures of skew and kurtosis are presented in Table 11.

3.4.1.3.2 Testing univariate assumptions of normality

Under the central limit theorem, in large samples (commonly defined as greater than 30 participants; Field, 2013) the sample distribution tends to be normal. This means that, regardless of the population distribution, parameter estimates of that population will have a normal distribution. The sample size of the current study far exceeds that specified by Field (2013) but, given the exploratory nature of this analysis, checks for univariate normality were conducted for each of the variables generated by the story stem assessment.

Boxplots, histograms, normality plots (P-P and Q-Q plots), and tests for skew and kurtosis were used to detect outliers and examine the distribution of story stem variables. Boxplots were generated to identify any outliers or extreme scores within the data. The number of outliers detected in the data were as follows: Aggressive themes: 3; Ostracising

themes: 2; Destructive themes: 1; Prosocial themes: 0; Narrative coherence: 2; Observed behavioural dysregulation: 2. No extreme scores were identified in the data. Each outlier was checked to ensure the case was not a coding or data entry error and all scores were found to be correct. All outliers were mild and judged to be reflective of normal variation in the data and thus were retained in the dataset. See Appendix I for the distributions of the story stem variables.

Through examining histograms for each of the story stem variables, the measures of aggressive and destructive themes looked slightly positively skewed (indicating a tendency for lower scores in the sample), with deviations from the diagonal line also seen on the P-P and Q-Q plots for these two variables. Skewness values for these variables were also a little above 1 (1.11 and 1.13). Although it is difficult to define exact ranges of skew and kurtosis that represent a normal distribution, these slightly elevated values of skewness fall within the standards set by George and Mallery (2019) of ±2. All values for kurtosis were below 1. The other measures derived from the use of the story stem measure (ostracising and prosocial themes, as well as measures of narrative coherence and observed behavioural dysregulation) appeared to be normally distributed with dots lining up along the diagonal line on the P-P and Q-Q plots, relatively symmetrical distributions on the histograms, and values of skew and kurtosis all below 1.

Kolmogorov-Smirnov and Shapiro-Wilk tests were also conducted to examine whether the distribution of scores deviated from a comparable normal distribution. The findings of these two tests were interpreted with caution given that conducting these tests with large samples can give a significant result (indicating a non-normal distribution) even for small and unimportant effects (Field, 2013). This seems to be the case here as all study variables had a significant result on these tests despite the fact most of the data looked normally distributed when the data was plotted.

Given the large sample size here, and the checks identifying what were judged to be only small deviations from normality, the data were assumed to be resistant to minor violations of normality and thus the data were not transformed.

Table 11

Descriptive Statistics for Study Variables Collected at the Two-Year Follow-Up

Variable[a]	*N*	*M*	*SD*	Median	Range	IQR	Skewness	Kurtosis
Story stem measure								
Aggressive themes	256	1.86	2.23	1.00	9.00	3.00	1.11	0.31
Ostracising themes	256	1.91	1.54	2.00	8.00	2.00	0.85	0.75
Destructive themes	256	1.23	1.33	1.00	6.00	2.00	1.13	0.76
Prosocial themes	256	3.68	1.93	4.00	8.00	3.00	-0.04	-0.74
Narrative coherence	256	10.42	2.78	11.00	13.00	4.00	-0.37	-0.24
Observed behavioural dysregulation	261	6.22	3.79	6.00	16.00	5.00	0.63	-0.25
Parent-reported child behaviour								
PPACS – total	261	23.35	8.82	22.00	48.00	11.00	0.61	0.33
PPACS – ADHD	261	9.64	5.37	9.00	27.00	6.00	0.78	0.31
PPACS – conduct	261	13.71	5.06	14.00	28.00	7.00	0.33	-0.05
CBCL – total	261	31.29	21.84	27.00	140.00	27.00	1.59	3.74
CBCL – externalising	261	11.96	8.74	10.00	44.00	12.00	1.14	1.39
SDQ – total	261	10.26	5.34	9.00	28.00	7.00	1.08	1.24
SDQ – externalising	261	6.97	3.70	7.00	18.00	5.00	0.73	0.25
SDQ – prosocial	261	7.54	1.85	8.00	9.00	3.00	-0.58	-0.13

Variable[a]	*N*	*M*	*SD*	Median	Range	IQR	Skewness	Kurtosis
Parenting practices and wellbeing								
Parental overreactive discipline	261	3.79	0.48	3.70	3.70	0.66	-0.43	1.35
Parental lax discipline	261	3.70	0.39	3.73	2.64	0.55	-0.27	1.00
GAD-7	260	4.03	4.32	3.00	20.00	4.00	1.65	2.55
PHQ-9	260	3.88	4.29	3.00	23.00	4.00	1.85	3.42
Teacher-reported child behaviour								
SDQ – total	194	6.95	5.58	6.00	28.00	7.25	1.11	1.41
SDQ – externalising	194	4.21	4.03	3.00	19.00	6.00	1.10	0.76
SDQ - prosocial	194	7.11	2.25	7.00	10.00	3.00	-0.53	-0.26

PPACS = Preschool Parental Account of Childhood Symptoms; CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire; GAD-7 = Generalised Anxiety Disorder-7; PHQ-9 = Patient Health Questionnaire-9

3.4.2 Aim 1: Examining the feasibility of delivering the measure

All children present at the two-year follow-up assessment were offered the opportunity to complete the story stem measure. A total of 256 children fully completed the measure, representing a completion rate of 94% for those families where measure administration was attempted. Reasons for non-completion are outlined in Table 12.

Table 12

Story Stem Battery Completion Status for Full HSHS Sample

Story stem battery completion status	*n* (%)
Measure completion attempted	
Story stem measure fully completed	256 (85)
Story stem measure partially completed	5 (2)
Child did not assent to measure completion	11 (4)
Measure completion not possible	
Visit conducted remotely – family moved abroad	7 (2)
Visit conducted remotely – family were unable to complete full home visit	7 (2)
Family did not complete any of their 24-month visit	14 (5)

N = 300

A total of 16 children did not complete or only partially completed the story stem battery when given the opportunity to do so. In order to explore whether there are certain factors that may increase or decrease the likelihood of a measure like this being completed by a child, some of the characteristics of these 16 children were explored.

By the time of the two-year follow-up assessment, 11 children (4%) had received a clinical diagnosis (e.g., autism, ADHD) from a qualified health or educational professional or were currently undergoing formal assessment through these services. Of these 11 children, six did not provide assent to complete any of the story stem measure. This gives a 45% completion rate in children currently under the care of CAMHS or equivalent services, compared to a completion rate of 96% in non-referred children. While this implies that assent rates for a measure like this may be lower in a more clinical population, a clinical diagnosis should not be taken to indicate that the children cannot complete the measure or are unlikely

to wish to do so, given that five children in this group did choose to complete the battery in full.

Five children only partially completed the measure. All of them completed the measure set-up (selecting dolls that represented themselves and their family members), the warm-up (a party with the dolls) and the first story stem (spilt juice scenario) before disengaging. In order to examine whether there were any common differences in children who disengaged compared to those who completed the measure in full, demographics and data of these five participants were considered. Four of the five children (80%) who disengaged were boys and their mean age was 40.6 months (*SD* = 5.9), which is similar to the mean age of the overall sample (40.0 months). During the portion of the measure that was completed by these children, the mean score assigned for behavioural dysregulation was 10.0 (*SD* = 5.7) which is higher than the mean of the children who completed the full battery (*M* = 6.2; *SD* = 3.7). However, 51 other children (20%) also scored 10 or above on the behavioural dysregulation measure so high levels of dysregulation did not prohibit the majority of children participating fully in the story stem battery.

3.4.2.1 Exploring children's responses to the story stems

To convey the kinds of data obtained through the story stem battery, some exemplar verbal responses to the stories are presented in Table 13. It should be noted that, in scoring, both verbal and non-verbal communication from the child was considered. Thus, providing a full and complete picture of the children's responses in text is a challenge. As children were given global scores for their responses to each story stem the exemplar data represents their response to the story prompt from start to finish.

Table 13

Illustrative Data Related to the Constructs Derived from the Story Stem Measure

Construct	Illustrative data extract
Aggressive (high); *Three's a Crowd*	Kick, kick. Emily kicked her away! And kicked everybody away. And then, **Mia** said "No, not funny" and then she left her out. And Dad said, "You're not allowed to do that, it's never good to do that, you'll have to go on time out". [Emily says] "I don't want to, so I'm just hitting you." [*Child repeatedly knocks Emily doll against Dad doll*] Now Daddy has blood on him, on his head. [*Mia doll hits against Emily doll repeatedly and says*] "Now, I'm stamping you dead, and now you're dead."… Then Emily said, "You're not allowed to [play with the ball]!" [Anna says] "Everybody's allowed to do it". [Emily says] "No, only me and **Mia**." [Anna says] "Everybody, see I'm **Mia's** and your friend, why are you not being my friend anymore?" **Mia** feels cross and sad because Emily is not being nice.
Prosocial; *Three's a Crowd*	The ball popped because it's a water ball. We're blowing it back up and we're putting Sellotape on it then it didn't go out. [Anna says] "Thank you, I want to play. Can I play too? We catched the ball all together" [*Researcher delivers second prompt about Emily not wanting to play with Anna*] [Emily says] I'm not going to play too [*Emily is placed in Mummy doll's lap*]. We're crying, we have no one else to play with.…[**Mia** feels] sad, both of them feel sad, because there's no more friends. [Emily feels] sad because she didn't want to play with Anna. That's not nice because sharing is caring, and you have to play together. [Emily says] "Sorry, can I play now?" [Anna says] "Yes, here's your ball, and now we can play football!" [Emily says] "Now me – [I kicked it] even farer!" [Mummy says] "I'm going to teach you how to play football, watch this. Like this!" Super far. [Anna says] "I can teach you – woah, I can do it."
Coherence (high); *Hurt Knee*	She gets up and tells Mummy and Daddy that her knee's hurting and then Mummy and Daddy bringed her to the hospital to get her knee fixed. And then she comes back and sits with Mummy and Daddy. And then Nora runs on the slide and then she slides on the slide and then she goes on the roundabout. Then, Daddy spins the roundabout fast and then he stops it because Nora wants to go on the small slide. **Mia** feels happy because her knee is almost better. … Nora falls over and bumps her head and then she sits down as well. [She's] sitting on Mummy's knee. Then Daddy goes on the roundabout and then the end.
Coherence (medium); *Hurt Knee*	It bleeds! Red makes bleed. And I say "Ow!" and then we talk about it to Mummy and Daddy. Look, [Mummy] is standing with him on her shoulders. See, they are flying! **Jack** can't fly. But now he flies. And Daddy comes in and he's flying – all of them are flying!...He runs very fast and he can't fall over – see. He goes to sleep. He hurt his knee and then he tells Mummy and then that's the end.
Coherence (low); *Hurt Knee*	Sam is okay. Sam falls over. Where's the slide? **Jack** is dancing, he's having a good dance. He can't fly now. Crash! And look, look, he's got another t-shirt. How's he got another t-shirt? They're dancing. He can't stand up, look, it's slippery – woah. Crash! They crashed together. Hey, look, my Dad's not on the side of the grass. Sam's okay, Sam's good for now. They fell over. There's no sun for Sam. He has a hat and then the sun goe'd away. Uh-oh, the sun. There's no sun.

Note: Bold text indicates study child. Pseudonyms used throughout. Jack and Mia represent the study children, Sam and Nora represent the friend in Hurt Knee, Emily and Anna represent friend 1 and 2 in Three's a Crowd.

3.4.2.2 Exploring the thematic content generated through the stems

Each story stem was assigned a score of between 0 and 3 for aggressive, ostracising, destructive, and prosocial themes. A total score for each of the four themes was generated by summing the scores children received across the three story stem scenarios. The mean scores of these themes are presented in Figure 4. Mean scores are also presented broken down by story stem in Figure 5.

We can see that overall children told stories with more prosocial themes (M = 3.68, SD = 1.93) compared to other themes. Aggressive themes were higher when the stories involved peers (story stem two and three) compared to family (story stem one) and highest when the child was presented with peer conflict in story stem three (M = 0.90, SD = 1.15). It is possible that the 'dilemma' presented by the researcher to the child for each story stem could have elicited certain themes. For instance, we can see in Figure 5 that destructive themes are higher in the first story stem, where the dilemma focused on the child spilling a jug of juice, and ostracising themes are highest in the third story stem, where the dilemma focused on one peer not wanting to play with another. However, exhibition of these behaviours was scored on a scale of 0 to 3 to pick up on variance in children's responses, even if the story beginning may have primed some children to continue these themes in their resolution of the story.

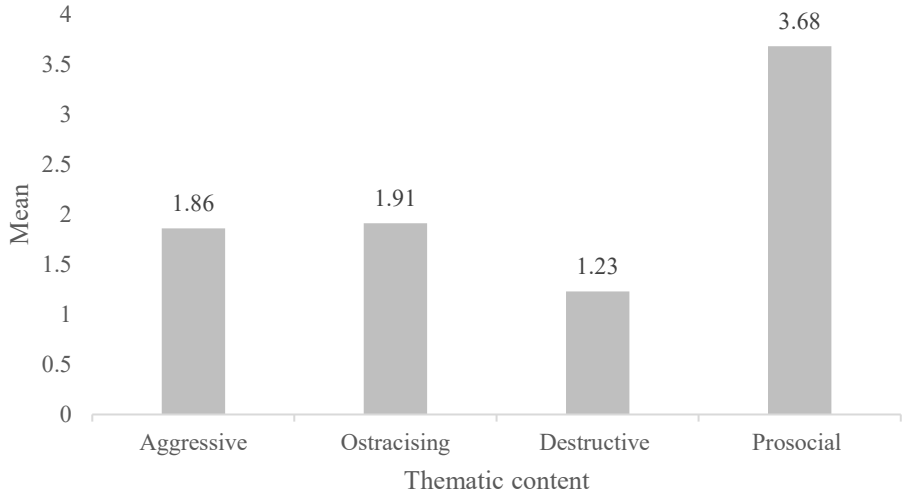


Figure 4. Mean scores of coded behaviours using summed scores across the three story stems.

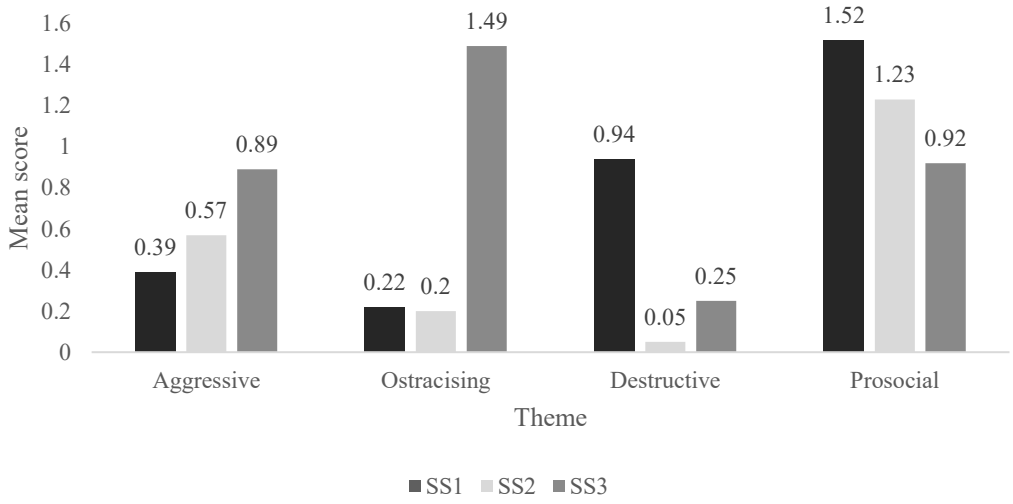


Figure 5. Mean scores of coded thematic content disaggregated by story stem.

3.4.2.3 Creating a composite variable of disruptive behaviours

The decision was made to create a composite theme of 'disruptive' behaviours exhibited in stories using the summed scores of aggressive, ostracising, and destructive themes. From a statistical perspective, combining these three themes into an overall score reduced the number of tests required. Also, based on their conceptual similarity, theoretically these behaviours may be tapping into conduct-related behaviours more broadly. A new variable of 'disruptive behaviours' was created (M = 5.00, SD = 4.10, range: 19). Its distribution was found to be slightly positively skewed and six outliers, but no extreme scores, were detected.

3.4.2.4 Exploring the internal consistency of scale scores

The final step in the initial exploration of the story stem data was to explore the internal consistency of the four scales. Cronbach's alpha (Cronbach, 1951) was used to test the relation between scores within the scales. Table 14 shows that disruptive themes, narrative coherence, and behavioural dysregulation had 'acceptable' to 'good' internal reliability but the expression of prosocial themes had 'poor' internal reliability (using rule of thumb acceptability guidelines; Brace, Kemp, & Snelgar, 2012). Cronbach's alpha is sensitive to the number of scale items (Cortina, 1993). This may explain why, despite having a higher mean inter-item correlation, prosocial story themes had a lower Cronbach's alpha than the disruptive story themes. Overall, this analysis revealed that the two scales adapted from pre-existing, well-established scales (narrative coherence and behavioural dysregulation) had greater reliability than the newly developed prosocial and disruptive scales.

Table 14

Internal Consistency of Story Stem Subscales

Scale	Number of items, *N*	Mean inter-item correlation, *r*	Cronbach's alpha, α
Prosocial story themes	3	.26	.51
Disruptive story themes	9	.23	.70
Narrative coherence	3	.56	.79
Behavioural dysregulation	6	.34	.75

3.4.3 Aim 2: Exploring concurrent associations between adult-reported child behaviour and constructs derived from children's responses to the story stems.

3.4.3.1 Exploring associations of story stem data with child behaviour

Correlational analyses were run to explore concurrent associations between children's responses to the story stems and primary caregiver-reported (PPACS, CBCL, and SDQ; Table 15) and teacher-reported (SDQ; Table 16) measures of child behaviour collected at the 24-month follow-up assessment. Given the exploratory nature of this analysis, both total scores and subscales (PPACS: conduct and ADHD subscales; CBCL: externalising; SDQ: externalising and prosocial) of the behaviour measures were included in analysis.

Before exploring associations between parent-, teacher-, and child-reported data, associations between the constructs derived from the story stem battery itself were examined. Table 15 shows there was a small, negative correlation between disruptive themes and prosocial themes (r = -.15, p = .02), a small, negative correlation between disruptive themes and narrative coherence (r = -.19, p = .002) and a medium, positive correlation between disruptive themes and observed behavioural dysregulation (r = .37, p < .001). There was a medium, negative correlation between prosocial themes and observed behavioural dysregulation (r = -.31, p = < .001), and a large, positive correlation between prosocial themes and narrative coherence (r = .68, p = < .001). These associations were in the expected direction, in that a child who told stories with a higher level of coherence and more prosocial themes was less likely to be dysregulated during storytelling or to include disruptive themes.

3.4.3.1.1 Story stem data and parent-rated child behaviour

Table 15 also shows that some small associations were observed between child-reported and parent-reported measures. Specifically, disruptive story themes were positively correlated with the total scores of the PPACS (r = .22, p < .001), CBCL (r = .24, p < .001), and the SDQ (r = .18, p = .004) with stronger associations on the measure's externalising subscales. Prosocial story themes (r = .21, p = .001) and children's narrative coherence (r = .14, p = .027) were positively correlated with the prosociality subscale of the SDQ but not with parent-rated reports of their child's attention or conduct difficulties. The observed measure of children's behavioural dysregulation during the story stem assessment was positively correlated with the total scores of the PPACS (r = .21, p < .001), CBCL (r = .15, p = .016), and the SDQ (r = .15, p = .014). The dysregulation measure was also negatively correlated with the prosociality subscale of the SDQ (r = -.21, p = .001).

Table 15

Bivariate Correlations between Story Stem Variables and Parent-Reported Child Behaviour

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Disruptive themes[a]	-											
2. Prosocial themes	-.15*	-										
3. Narrative coherence	-.19**	.68**	-									
4. Observed dysregulation	.37**	-.31**	-.36**	-								
5. PPACS - Total	.22**	.04	-.00	.21**	-							
6. PPACS - ADHD	.15*	.01	.01	.14*	.86**	-						
7. PPACS - Conduct	.22**	.06	-.01	.22**	.84**	.43**	-					
8.CBCL - Total	.24**	-.02	-.05	.15*	.59**	.45**	.55**	-				
9.CBCL - Externalising	.23**	-.02	-.03	.20**	.62**	.46**	.59**	.90**	-			
10. SDQ - Total	.18**	-.03	-.05	.15*	.62**	.52**	.53**	.83**	.79**	-		
11. SDQ - Externalising	.18**	.01	.01	.19**	.62**	.53**	.53**	.71**	.83**	.87**	-	
12. SDQ - Prosocial	-.07	.21**	.14*	-.21**	-.28**	-.18**	-.30**	-.40**	-.47**	-.42**	-.39**	-

Note. * $p < .05$, ** $p < .01$

a Composite score of aggressive, ostracising, and destructive themes

PPACS = Preschool Parental Account of Childhood Symptoms; CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire

3.4.3.1.2 Story stem data and teacher-rated child behaviour

Table 16 shows that some small associations were observed between child-reported and teacher-reported data. Children's expression of prosocial story themes was positively associated with teacher-rated prosociality (r = .24, p = .001) and negatively associated with total behaviour problems (r = -.20, p = .005). Children's narrative coherence was positively associated with teacher ratings of prosocial behaviour (r = .24, p = .001) and negatively associated with total behaviour problems (r = -.22, p = .002). The observed measure of children's behavioural dysregulation during the story stem assessment was positively correlated with total behaviour problems (r = .21, p = .003) and negatively associated with the prosocial subscale (r = -.25, p < .001) on the teacher-rated questionnaire. However, unlike the parent-reported measures of children's behaviour problems, teacher-rated behaviour problems were not associated with children's expression of disruptive story themes (r = 0.08, p = .26).

Table 16

Bivariate Correlations between Story Stem Variables and Teacher-Reported Child Behaviour

Variable	1	2	3	4	5	6	7
1. Disruptive themes	-						
2. Prosocial themes	-.15*	-					
3. Narrative coherence	-.19**	.68**	-				
4. Behavioural dysregulation	.37**	-.31**	-.36**	-			
5. SDQ - Total	.08	-.20**	-.22**	.21**	-		
6. SDQ - Externalising	.11	-.16*	-.17*	.26**	.86**	-	
7. SDQ - Prosocial	-.10	.24**	.24**	-.25**	-.60**	-.57**	-

Note. * p < .05, ** p < .01
SDQ = Strengths and Difficulties Questionnaire

3.4.4 Aim 3: Explore whether child behaviour, measured through the story stem assessment, varied based on characteristics that typically predict child behaviour

This section considers the relationships between the constructs obtained through the story stem assessment battery and factors that are often related to children's behaviour. I first explored associations between story stem domains (thematic content and coherence of the stories, as well as the child's behavioural dysregulation) and factors external to the child (parental psychopathology, parental educational attainment, and parenting practices). Next, I examined any effects of child gender and age on their responses and behaviour during the completion of the story stem battery.

3.4.4.1 Exploring associations between story stem data with parent-related variables

Table 17 shows associations between the story stem variables (disruptive themes, prosocial themes, narrative coherence, and behavioural dysregulation) and parental psychopathology (PHQ9 and GAD7), parental discipline practices (Parenting Scale) and parental educational attainment. To provide more context to these findings, I also explored relationships between the same variables (parental psychopathology, discipline practices, and educational attainment) but with parent-reports of child behaviour, using the PPACS interview and CBCL and SDQ questionnaires (see Table 18) and teacher-reports of child behaviour using the SDQ (see Appendix J, Table 53). Each variable is now considered in turn, with the associations with child-, parent-, and teacher-rated child outcomes discussed together.

Parental psychopathology. Parental psychopathology was assessed using measures of parental low mood (PHQ-9) and anxiety (GAD-7). There was a small positive association between disruptive themes in the children's story stems and parental low mood (r = .19, p = .002) and anxiety (r = .17, p = .005). There was less evidence of associations between observed behavioural dysregulation and parental anxiety (r = .10, p = .12) and mood (r = .10, p = .13). Prosocial themes and narrative coherence during the stories did not appear to be associated with parental psychopathology (see Table 17). Comparatively, children's behaviour as rated by their primary caregiver was moderately to strongly positively correlated with parental low mood and anxiety across the PPACS, CBCL and SDQ measures (rs = .35-.52, p < .001). There were also small negative associations between children's prosociality, as rated by their primary caregiver, and parental low mood (r = -.21, p = .001) and anxiety (r = -.19, p = .002) (see Table 18). However, no statistically significant associations were found

between child behaviour as rated by their teachers and either measure of parental psychopathology (*r*s -.04-.11; see Appendix J).

Parenting practices. Parenting practices were assessed using the parenting scale which measures caregiver use of overreactive and lax discipline. There were no statistically significant associations between children's story stem themes, coherence, or behavioural dysregulation, and their primary caregiver's use of dysfunctional discipline (*r*s .00-.08; see Table 17). Similarly, there also appeared to be no associations between parent-reported child behaviour and parental discipline practices (*r*s .02-.06; See Table 18).

Parental educational attainment. There was a small, negative association between children's disruptive story stem themes and parental education ($r = -.13$, $p = .029$) (see Table 17). Children whose caregivers had a higher level of educational attainment were less likely to express disruptive themes. Parental education was not significantly associated with children's prosocial themes, their coherence, or their dysregulation. Comparatively, there were small to moderate negative associations between parent-reported measures of children's behaviour and parental education (*r*s -.18, -.24, -.34) but no such associations with parent-reported prosociality ($r = .09$). A similar pattern of results were found for teacher-rated behaviour (total problems: $r = -.19$; externalising problems: $r = -.16$; prosociality: $r = .13$; Appendix J).

Table 17

Bivariate Correlations between Story Stem Variables and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment

Variable	1	2	3	4	5	6	7	8	9
1. Disruptive themes	-								
2. Prosocial themes	-.15*	-							
3. Narrative coherence	-.19**	.68**	-						
4. Behavioural dysregulation	.37**	-.31**	-.36**	-					
5. Parent depressive symptoms	.19**	.05	.09	.10	-				
6. Parent anxiety symptoms	.17**	.04	.07	.10	.78**	-			
7. Overreactive discipline	.01	.01	.01	.04	.07	-.00	-		
8. Lax discipline	-.02	.06	.08	.04	-.05	-.04	-.02	-	
9. Parent education	-.13*	-.08	.05	-.09	-.14*	-.13*	-.15*	.05	-

Note. * $p < .05$, ** $p < .01$

Table 18

Bivariate Correlations between Parent-Rated Child Behaviour and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment

Variable	1	2	3	4	5	6	7	8	9
1. Parent-rated PPACS	-								
2. Parent-rated CBCL	.59**	-							
3. Parent-rated SDQ	.62**	.83**	-						
4. Parent rated SDQ prosocial	-.28**	-.40**	-.42**	-					
5. Parental depressive symptoms	.35**	.52**	.50**	-.21**	-				
6. Parental anxiety symptoms	.25**	.36**	.34**	-.19**	.78**	-			
7. Parental overreactive discipline	.06	-.01	-.04	.06	.07	-.00	-		
8. Parental lax discipline	.05	.02	-.05	-.04	-.05	-.04	-.02	-	
9. Parental educational attainment	-.34**	-.18**	-.24**	.09	-.14*	-.13*	-.15*	.05	-

Note. * $p < .05$, ** $p < .01$

CBCL = Child Behavior Checklist; PPACS = Preschool Parental Account of Childhood Symptoms; SDQ = Strengths and Difficulties Questionnaire

3.4.4.2 Exploring gender differences in the story stem data

Independent *t*-tests were used to explore gender differences in the story stem data. There was a small difference in sample size for boys (n = 136) and girls (n = 120). As this difference was small no further action was taken. The distributional assumptions necessary for an independent *t*-test have already been discussed in section 3.4.1.3. Levene's test, a test of the homogeneity of variance, was also run although interpretation of this test in large samples should be treated with caution (Zimmerman, 2004). Standardised effect sizes (Cohen's *d*) were computed using the difference in means between the two groups and dividing this by the pooled standard deviation. Cohen's *d* effect sizes can be defined as 0.20 = small, 0.50 = medium, and 0.80 = large (Cohen, 1969, 1988). See Figures 20-23 in Appendix K for the distribution of narrative dimensions by gender.

Disruptive themes. For disruptive themes, Levene's test was significant (F = 8.40, p = .004) indicating unequal variances. Consequently, the outcome of the *t*-test for 'equal variances not assumed' was used for interpretation. On average, male participants scored more highly for disruptive themes (M = 6.18, SE = 0.37) than female participants (M = 3.68, SE = 0.31). This difference, -2.50, 95% CI [-3.45, -1.55], $t(251) = -5.18$, $p < .001$, represented a medium-sized effect, Cohen's d = 0.64.

Prosocial themes. For prosocial themes, Levene's test was not significant (F = 0.49, p = .48) indicating equal variance. On average, female participants scored more highly on prosocial themes (M = 4.35, SE = 0.17), than male participants (M = 3.08, SE = 0.15). This difference, 1.27, 95% CI [0.82-1.72], $t(254) = 5.56$, $p < .001$, represented a medium-sized effect, Cohen's d = 0.70.

Narrative coherence. For narrative coherence, Levene's test was not significant (F = 0.26, p = .61) indicating equal variance. On average, female participants scored more highly for narrative coherence (M = 11.41, SE = 0.23), than male participants (M = 9.55, SE = 0.23). This difference, 1.86, 95% CI [1.21, 2.50], $t(254) = 5.65$, $p < .001$, represented a medium-sized effect, Cohen's d = 0.71.

Observed behavioural dysregulation. For observed behavioural dysregulation, Levene's test was significant (F = 16.47, $p < .001$) indicating unequal variances. Consequently, the outcome of the *t*-test for 'equal variances not assumed' was used for interpretation. On average, male participants (M = 7.43, SE = 0.34) scored more highly for behavioural dysregulation during completion of the story stem battery than female

participants (M = 4.70, SE = 0.26). This difference, -2.73, 95% CI [-3.57, -1.89], t(247) = -6.38, p < .001, represented a large effect size, Cohen's d = 0.79.

3.4.4.3 Exploring adult-reported gender differences in child behaviour

Given the consistent gender differences in both thematic content, narrative coherence, and observed behavioural dysregulation derived through the story stem analysis, parent- and teacher-reported gender differences of child behaviour were also examined in post-hoc analysis. The purpose of this was to examine if comparable and consistent gender differences were seen in adult-reported measures as had been seen in the child-reported measure. This data is presented in Table 19.

Parent-rated child behaviour. In the parent-rated data there was a trend of boys receiving higher scores for problem behaviours and lower scores for prosociality than girls. However, the *t*-tests exploring these gender differences found only very small effects for the total scores of the PPACS (d = 0.09), CBCL (d = 0.14) and SDQ (d = 0.12), and the conduct subscale of the PPACS (d = 0.12). For the CBCL externalising subscale, on average, male participants (M = 13.09, SE = 0.76) scored higher than female participants (M = 11.04, SE = 0.77). This difference, -2.05, 95% CI [-4.19, 0.93], t(254) = -1.88, p = 0.06, represented a small effect size, Cohen's d = 0.24. For the SDQ prosocial subscale, on average, male participants (M = 7.24, SE = 0.17) scored lower than female participants (M = 7.84, SE = 0.15). This difference, 0.60, 95% CI [0.34, 2.63], t(254) = 2.63, p = 0.01, represented a small effect size, Cohen's d = 0.34.

Teacher-rated child behaviour. The teacher-rated data revealed gender differences that were more consistent in magnitude with those seen in the children's story stem data. For the SDQ externalising subscale, on average, male participants (M = 5.12, SE = 0.39) scored higher than female participants (M = 3.04, SE = 0.40). This difference, -2.09, 95% CI [0.56, -3.20], t(192) = -3.70, p = .000, represented a medium effect size, Cohen's d = 0.54. For the SDQ prosocial subscale, on average, male participants (M = 6.66, SE = 0.22) scored lower than female participants (M = 7.68, SE = 0.22). This difference, 1.02, 95% CI [0.32, 0.39], t(192) = 3.20, p = .002, represented a medium effect size, Cohen's d = 0.47.

Table 19

Parent- and Teacher-Reported Child Behaviour by Child Gender

Outcome	Female mean (SD)	Male mean (SD)	Mean difference (95% CI)	Effect size[a]	*t*-value	*p*-value
Parent-rated behaviour[b]						
PPACS – Total	23.05 (9.34)	23.82 (8.40)	-0.76 (-2.95, 1.42)	0.09	-0.69	0.49
PPACS – Conduct	13.46 (5.11)	14.07 (5.01)	-0.62 (-1.86, 0.63)	0.12	-0.97	0.33
CBCL – Total	30.04 (21.25)	33.09 (22.31)	-3.05 (-8.43, 2.33)	0.14	-1.12	0.27
CBCL – Externalising	11.04 (8.42)	13.09 (8.91)	-2.05 (-4.19, 0.93)	0.24	-1.88	0.06
SDQ – Total	10.02 (5.56)	10.68 (5.11)	-0.66 (-1.98, 0.65)	0.12	-1.00	0.32
SDQ – Prosocial	7.84 (1.67)	7.24 (1.97)	0.60 (0.15, 1.05)	0.34	2.61	0.01
Teacher-rated behaviour[c]						
SDQ – Total	5.45 (5.03)	8.12 (5.73)	-2.67 (-4.22, -1.11)	0.49	-3.39	.001
SDQ – Externalising	3.04 (3.69)	5.12 (4.06)	-2.09 (-3.20, -0.97)	0.54	-3.70	.000
SDQ – Prosocial	7.68 (2.05)	6.66 (2.30)	1.02 (0.39, 1.64)	0.47	3.20	.002

Note. PPACS = Preschool Parental Account of Childhood Symptoms; CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire
a Effect size is Cohen's *d;* b $N = 256$; c $N = 194$

3.4.4.4 Exploring age differences in the story stem data

Effects of child age at the time of measure completion on outcomes derived from the story stems were also examined. Correlational analysis presented in Table 20 shows that prosocial themes ($r = .33$, $p = <.001$) and narrative coherence ($r = .31$, $p <.001$) were significantly positively correlated with child age. Disruptive themes and behavioural dysregulation were not associated with age. See Figures 24 and 25 in Appendix L for the story stem data plotted by age.

Table 20

Bivariate Correlations between Child Age and Story Stem Constructs

Variable	1	2	3	4	5
1. Age	-				
2. Disruptive themes	.09	-			
3. Prosocial themes	.33**	-.15*	-		
4. Narrative coherence	.31**	-.19**	.68**	-	
5. Observed dysregulation	-.07	.37**	-.31**	-.36**	-

Note. * $p < .05$, ** $p < .01$

3.4.4.5 Exploring age by gender differences in narrative coherence

Finally, given the previous literature on differences in the developmental trajectories of narrative coherence for boys and girls, the gender and age differences in this construct were explored together. Figure 6 shows children's narrative coherence broken down by gender and age in three-monthly increments. We can see that at age three boys and girls follow a similar developmental trajectory in terms of improvements in narrative coherence but with boys consistently lagging behind girls. Between the ages of four and five though, girls' narrative coherence stabilised while boys appeared to continue 'catching up' with the girls.

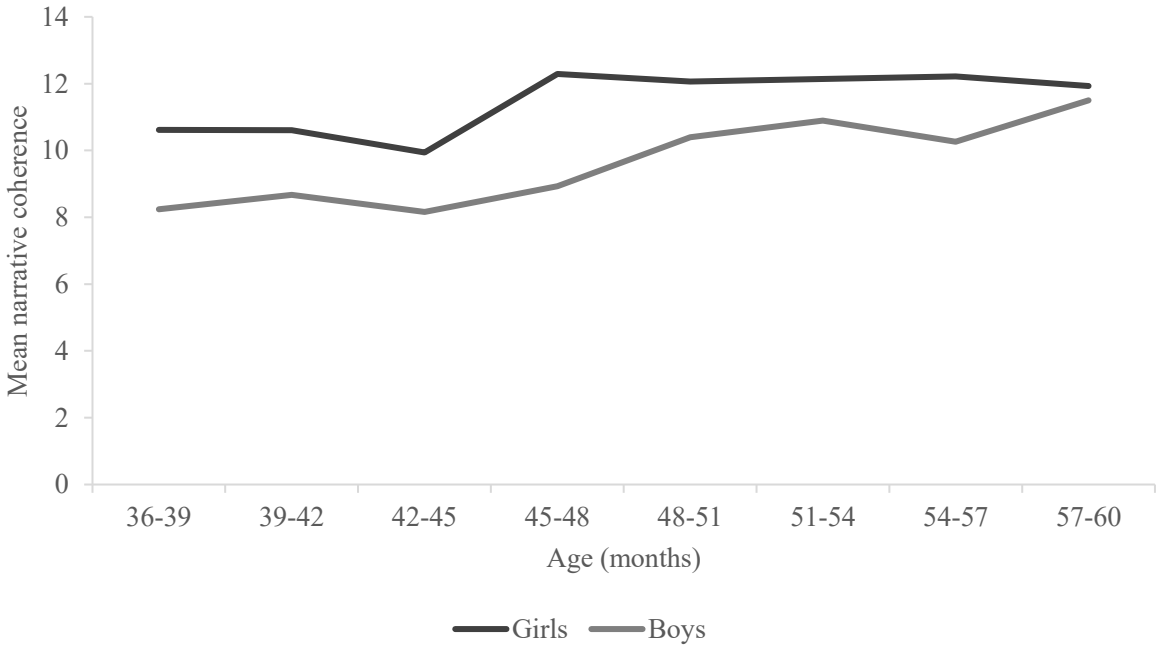


Figure 6. Children's narrative coherence disaggregated by age and gender.

3.5 Discussion

This study explored the use of a brief, dolls house play story stem procedure in a sample of three- and four-year-olds recruited based on early reports of behaviour problems. The measure was acceptable to children and families and feasible to deliver as part of a wider assessment. When given the opportunity to, 94% of children fully completed the story stem battery (another 2% partially completed the measure while 4% did not provide assent). Specific aspects of the themes and structure of children's play narratives elicited through this developmentally sensitive tool were found to be associated with both parent- and teacher-reports of child behaviour in the directions that would be expected. Further to this, factors known to predict child behaviour problems, including child gender, parental education level, and parental depression, were also associated with story stem constructs in the ways we would expect. Here, the research findings are discussed in detail, before the strengths, limitations, and implications of the study are considered.

3.5.1 Summary of findings

3.5.1.1 Children's story stem outcomes and adult-rated child behaviour

Overall, there were consistent patterns of association between story stem domains and the anticipated constructs of the concurrently recorded parent- and teacher-rated data in the expected directions. Disruptive story themes were positively associated with three measures of parent-reported behaviour problems, narrative coherence and prosocial story themes were positively correlated with parent- and teacher-reported prosociality, while observed behavioural dysregulation during task completion was positively associated with behaviour problems and negatively associated with prosociality rated by children's parents and teachers. These findings are now discussed in more detail in the context of existing literature.

3.5.1.1.1 Disruptive behaviours

Children's expressions of disruptive story themes were positively associated with concurrent parent-reported behaviour problems measured using a researcher-led interview and two questionnaires. This might suggest that children's play behaviours reflect tendencies to respond to real-life challenges in a disruptive way (Crick & Dodge, 1994). This finding is consistent with explorations of the narratives of non-clinical four- and five-year-olds (von Klitzing et al., 2000; Warren et al., 1996), and extends our insight to three-year-olds.

3.5.1.1.2 Prosocial themes

Children's expressions of prosocial story themes were concurrently associated with parent- and teacher-reported prosociality and negatively associated with teacher-rated behaviour problems. Adult-reported prosociality focused on exhibitions of empathy, sharing, helpfulness, and kindness. Variation in the measurement of 'prosocial behaviour' limits direct comparisons to prior studies. However, previously children's prosocial narrative themes have been associated with their social competence (Laible, Carlo, Torquati, & Ontai, 2004), emotional coherence (Oppenheim, Nir, Warren, & Emde, 1997), cheating behaviours (Woolgar et al., 2001), and moral conduct and conscience (Kochanska, Padavich, & Koenig, 1996). The current findings extend our understanding to a large sample of children as young as three years old, something investigated in only a minority of small-sampled studies previously (e.g., Oppenheim, Emde, Hasson, & Warren, 1997; $N = 51$; 3.5-5.5 years).

3.5.1.1.3 Narrative coherence

A coherent narrative structure (i.e., the degree to which children addressed and resolved the story conflict using a storyline that made sense; Robinson, 2007) was associated with parent- and teacher-reported prosociality and children's expression of prosocial story themes. It has been theorised that narrative coherence is a feature of the attachment system (Green et al., 2007). Associations between secure attachment classification and/or sensitive parenting and prosocial behaviours have been reported previously (e.g., Laible, 2007; Murphy & Laible, 2013; Paulus et al., 2015), although this finding has not been consistent across toddler-aged samples (e.g., Loheide-Niesmann, de Lijster, Hall, van Bakel, & Cima, 2021). With this in mind, associations between narrative coherence and prosociality in this study may to some extent reflect aspects of children's attachment representations. The negative association between narrative coherence and teacher ratings of behaviour problem is consistent with previous studies of school-aged children exhibiting clinically significant difficulties (e.g., Lorch, Milich, Flake, Ohlendorf, & Little, 2010; Scholtens, Rydell, Bohlin, & Thorell, 2014; Zenaro, Rossi, de Souza, & Giacheti, 2019), although no such association was found for parent-rated behaviour. A study from Oppenheim, Emde, Hasson, et al. (1997) offers possible insight into the high degree of association between children's narrative coherence and expression of prosocial themes. In their study, they found that most three-year-olds who were able to acknowledge and respond to the story dilemma did so in a prosocial way. However, this association was much weaker in five-year-olds, suggesting that the features of children's

acknowledgement and resolution become more differentiated and independent with age. Perhaps the high degree of association is reflective of the developmental stage of this sample.

3.5.1.1.4 Behavioural dysregulation

Children's behavioural dysregulation during the story stem task was negatively associated with their narrative coherence, positively associated with behaviour problems, and negatively associated with prosociality rated by their parents and teachers. The story stem activity mirrors behavioural expectations commonly seen in the classroom (e.g., remaining seated, following instructions). Children's ability to regulate their behaviour during this task may therefore be reflective of behavioural difficulties children can experience as they transition into formal education (Smith-Donald et al., 2007; Willoughby, Kupersmidt, Voegler-Lee, & Bryant, 2011). The slightly stronger association seen with teacher-rated difficulties could be because these behaviours are more noticeable in educational settings.

Collectively, these findings suggest that aspects of children's narrative representations relate to their observable behaviours rated by the adults around them. The magnitude of these associations (*r*s .14-.26) are consistent with previous findings in cross-informant child psychology research, with a meta-analysis of 269 samples revealing parent, teacher, and child agreement ranging from .20 to .27 (Achenbach et al., 1987). Adult-child agreement on measures of internalising and externalising difficulties (*r*s .08-.19; Ringoot et al., 2013); preschool experiences (*r*s .11-.24; Ruzek et al., 2020); and happiness in school (*r*s .06-.21; Allen et al., 2018) are also comparable to the current findings. The modest but consistent associations can be seen as supportive of the value of undertaking multi-informant approaches, with the view that each informant can offer unique insight into the child's behaviour (Arseneault et al., 2005).

It is worth considering though that the modest levels of agreement may also be reflective of weaknesses in the story stem measure. Administering and analysing measures that cater to young children's short attention span, emerging communication skills, and cognitive abilities is an inherent challenge in this field (Luby et al., 2007). Alternatively, children's responses could reflect less embedded symptoms, more transient experiences, or issues that are salient to children themselves but do not map onto what adults prioritise. Children were transitioning across childcare and educational settings at the time of the assessment and may have been providing us with unique insight into their experiences of these new contexts (O'Farrelly,

Booth, Tatlow-Golden, & Barker, 2020). Finally, differences in the focus, domains, and items of adult- and child-reported measures may have contributed to discrepancies in reporting.

3.5.1.2 Children's story stem outcomes and child gender and age

Clear gender differences were identified in children's stories. As a group, boys told less coherent, more disruptive, and less prosocial stories, while exhibiting more dysregulated behaviour. This finding is consistent with previous reports of more coherence and prosociality in girls' narratives (Murray et al., 1999; Zahn-Waxler et al., 1994; Zahn-Waxler et al., 1996) and more aggressive and incoherent patterning in boys' narratives (Page & Bretherton, 2003; von Klitzing et al., 2000) in both clinical and non-clinical populations. Possible explanations for these findings are numerous, but several are now considered.

First, these differences may be reflective of gender differences in children's mental health. Behaviour problems are more prevalent in boys and internalising difficulties are more prevalent in girls (conduct disorder: 69% male; ADHD: 86% male; anxiety: 59% female; depression: 57% female; Green et al., 2005). Woolgar et al. (2001) notes that a high occurrence of prosocial themes may be reflective of an empathic over-involvement and sense of responsibility for others' problems as a pathway to internalising difficulties. Thus, disruptive and dysregulated behaviours during story stem completion from boys and a tendency of girls to include prosocial themes may be linked to these gender differences in mental health. However, this is unlikely to fully explain these observations given that gender effects in parent-reported (although importantly not teacher-reported) data were much smaller.

Second, findings could represent gender differences in play behaviours (e.g., Dunn & Hughes, 2001; Rao, Fink, & Gibson, 2021; Zahn-Waxler et al., 2008). Boys have been found to exhibit aggressive behaviours during free play that they can inhibit during other activities (Ostrov & Keating, 2004). This could help to explain why clear gender differences were found in a play-based measure but not through a parent-reported questionnaire measure of general behaviour. Furthermore, girls may have been able to more readily engage with the pretend play nature of the task, reflected in their greater coherence, as parents have been found to engage in more pretend play with girls than boys (Lindsey & Mize, 2001).

Third, findings may reflect the gender-specific language acquisition gap. Girls outperform boys on word comprehension (Fenson et al., 1994), vocabulary production (Feldman et al., 2000; Galsworthy, Dionne, Dale, & Plomin, 2000), sentence complexity, and use of expressive language (Bornstein, Hahn, & Haynes, 2004) from eight months until about

six years of age, when this advantage begins to fade (Beltz et al., 2013). This proposal is supported by the data in the current study, where boys were found to consistently lag behind girls in their narrative coherence at age three to four, with signs that they were 'catching up' with girls between the ages of four and five. Parents have also been found to engage in gendered interactions, talking about a wider variety of emotions and engaging in more literacy-related activities with their daughters than sons (Leavell, Tamis-LeMonda, Ruble, Zosuls, & Cabrera, 2012; Peterson & McCabe, 2004). Coherent storytelling and the articulation of prosocial themes is more reliant on children's language and understanding of emotions than aggressive or destructive themes, which could more easily be communicated non-verbally (e.g., banging the dolls together to demonstrate hitting).

Fourth, findings may reflect the socialisation of girls to express empathetic feelings and suppress 'bad' behaviours (Brody, 1993). Girls have been found to be more likely to 'mask' socially undesirable behaviours in interactions outside of the home (Tomlinson, Bond, & Hebron, 2020; Mandy et al., 2012), and more likely to engage with the researcher's prompts and questions in story stem batteries (Oppenheim, Nir, et al., 1997). Taken together, these tendencies may have contributed to girls more coherent and prosocial behaviour during task completion with the researcher. The effects of 'masking' could also extend to educational settings, supported by the larger gender differences in teacher-rated child behaviour seen in this sample, with lower teacher reports of behaviour problems and higher prosociality for girls. One possibility is that both parent-reported data (where gender differences were relatively small) and observed and teacher-rated behaviour (where gender differences were larger) may both represent accurate depictions of child behaviour but reflect differences in children's conduct across contexts and settings.

Finally, child age was associated with narrative coherence and expression of prosocial themes. Steep developmental increases in children's abilities to address and resolve dilemmas in a prosocial and coherent way through the preschool years have been reported (Masashi & Tsutomu, 2012; Oppenheim, Nir, et al., 1997; Warren et al., 1996). Oppenheim, Emde, Hasson, et al. (1997) found that, in one story, only 45% of three-year-olds acknowledged the dilemma compared to 100% of five-year-olds. The associations with age may also reflect the dependence of coherent and prosocial narratives on language abilities, something that typically increases with age (Dealy, Mudrick, & Robinson, 2019; Fiorentino & Howe, 2004).

3.5.1.3 Children's story stem outcomes and parental psychopathology, education, and discipline style

No associations were found between parental mood, anxiety, and education and children's prosocial themes, narrative coherence, or observed behavioural dysregulation. However, parents' education level and depressive and anxiety symptoms were associated with children's expression of disruptive story themes. These associations were smaller than those seen when child behaviour was rated by parents though. One explanation for this is that parents' interpretation and reporting of their child's behaviour may be affected by their own mental health, a bias that has been widely debated (e.g., De Los Reyes & Kazdin, 2005; Najman et al., 2001). Some clear evidence was provided by Ringoot et al. (2015) though using a puppet-interview child-report measure. Estimates were twice as large in instances where parents rated both their own psychopathology and their child's behaviour problems. The findings of the current study follow a similar pattern to this. The possibility of parental bias in reporting is further supported by the fact there was no evidence of association between teacher-rated child behaviour and parental psychopathology in this study.

Parental discipline style was not associated with any components of children's narratives. Perhaps other more pervasive features of parent-child interactions such as parental sensitivity or communication style (e.g., 'elaborative reminiscing'; Bergen, Salmon, Dadds, & Allen, 2009) have a greater effect. Even so, in the preschool years discipline is a common feature of caregiver-child interaction (Power & Chapieski, 1986) and harsh discipline styles have previously been associated with children's self-reported behaviour (e.g., Mackenbach et al., 2014). Given this prior research, coupled with the fact that there were also no associations between parent-reported child behaviour and discipline, the utility of the parental discipline measure may require further investigation. For instance, consistency in ratings over time or comparisons to observed measures of discipline (interactions that were available in HSHS but that the coding of which went beyond the scope of this thesis) could provide a clearer idea of whether measurement issues may partially explain the lack of associations seen here.

3.5.2 Strengths and limitations

Here, I first consider the strengths and limitations of the sample and design and then the measurement of variables.

3.5.2.1 Sample and design

This section focuses on three strengths of the study's sample and design: its inclusion of children's own perspectives, its focus on children considered to be at an increased likelihood of enduring behaviour problems, and its comprehensive recruitment strategy. Limitations are then considered, with a focus on the extent to which we can generalise these findings to non-clinical, highly clinical, and international populations, and a consideration of some of the constraints of the study's design.

3.5.2.1.1 Inclusion of children's own perspectives

Inviting three- and four-year-olds to contribute to our understanding of their development is a key strength of this study. The need and value of integrating children's perspectives into matters that affect their lives was mandated over three decades ago (United Nations Committee on the Rights of the Child, 2005; United Nations Convention on the Rights of the Child, 1989). To date though, such efforts have almost exclusively included school-aged children and predominantly those in middle childhood (aged 7+) and adolescence (Pascal & Bertram, 2009; Urbina-Garcia, 2019). The findings of the current study benefit from a larger sample than most studies in the field, and make a unique contribution to our understanding of the ways very young children, as young as three years old, can provide meaningful accounts of their experiences when given authentic and appropriate opportunities to do so (O'Farrelly, 2021). This work must continue and advance to ensure children are heard, starting in their earliest years.

3.5.2.1.2 A focus on an under-studied population

Disparities in informants' reporting can be higher in clinical samples compared to community samples (De Los Reyes & Kazdin, 2005; Kolko & Kazdin, 1993). Because of this, exploring the suitability of child-reported measures as part of a multi-informant approach may be especially valuable for these populations. The high level of feasibility of delivering the story stem measure with a sample of three- and four-year-olds recruited based on early signs of behaviour problems is encouraging and advances our understanding of the populations in which child-reported measures can be utilised.

3.5.2.1.3 Comprehensive recruitment strategy

Recruitment into clinical trials typically favours a homogenous group of middle-class, highly-educated families, leading to the under-representation of certain groups, and

threatening the external validity and generalisability of findings (Jordan et al., 2013). For the current study, recruitment was mostly conducted through NHS health visiting developmental reviews, a universal service for all families of one- and two-year-olds. This approach provided access to a group of families who were broadly representative of the communities they lived in in terms of ethnicity, employment status, and age. It should be noted though that the current sample were more highly educated than the general population, a common finding in randomised controlled trials involving children (Robinson, Adair, Coffey, Harris, & Burnside, 2016), and thus generalisability may still be limited.

3.5.2.1.4 Generalisability to other populations

Focusing on children considered to be more likely to develop enduring behaviour problems has been noted as a strength of this study. However, this could also limit the generalisability of findings. Effects found in a general population might be attenuated in this sample, as there may be less variability in the behaviour of participating children. Also, while the current findings indicate that the story stem measure is suitable for children exhibiting challenging behaviours, further investigations are needed to fully understand their suitability in groups experiencing highly significant, diagnosed clinical difficulties. Finally, the relevance of these findings to international contexts is unknown. The majority of the representation of children's voices in developmental research comes from North America, Western Europe, and Australia. Dolls play methodologies have been used in small samples of children in Ghana (Hosny, Danquah, Berry, & Wan, 2020), South Africa (Pritchett, Rochat, Tomlinson, & Minnis, 2013), Uganda (Dent & Goodman, 2020), and with families of refugee backgrounds from the Middle East (Dalgaard, Todd, Daniel, & Montgomery, 2016). However, further study across global populations is required. Key to this is establishing the cross-cultural relevance and sensitivity of story scenarios, props, and coding frameworks to safeguard again irrelevant or inappropriate conclusions being drawn by viewing children's data solely through a westernised lens.

3.5.2.1.5 Cross-sectional design

Finally, the cross-sectional design of this study generated correlational findings, limiting the inferences that can be made regarding causality or direction of effects. This also removed the possibility of exploring the test-retest reliability of the story stem measure or the stability of the children's representations over time. Given the brevity of the measure and the unique nature of the sample, these properties would have been particularly interesting to explore.

3.5.2.2 Measurement of study variables

Next, I discuss the strengths and limitations of the measurement of variables, focusing first on the story stem measure and then the adult-reported data. I consider the strengths of including children's perspectives before outlining limitations focused on: the novelty of aspects of the measure, the high level of resource required, issues around child assent and respondent validation, and the requirement to administer the measure face-to-face. For the adult-reported data, I propose that the multi-informant approach is a key strength of the study but reflect upon the possible influences of informant bias.

3.5.2.2.1 Story stem measure

3.5.2.2.1.1 Access to children's voices

The insight gained through administering the story stem measure with three- and four-year-olds lends support to the ability of brief, play-based measures to gather rich data from what are often regarded as 'difficult-to-assess' populations. Very young children form a group in which more formal interview methods tend to fail. The story stem batteries' use of play and storytelling, familiar concepts that often spontaneously interest young children, as well as tangible props, cater to children's developing verbal abilities and cognitive functioning, possibly generating a higher level of ecological validity. Such approaches may help to bridge the gap between behavioural observation and direct verbal interview methods (Bettmann & Lundahl, 2007; Plokar, Bisaillon, & Terradas, 2018).

3.5.2.2.1.2 Novel elements of the coding approach

This research was informed by two well-validated approaches to narrative assessment (the MacArthur Story Stem Battery and Dolls House Play method). However, features of the administration and coding approach were novel. First, in an effort to reduce participant burden and fatigue, data were drawn from only three story stem scenarios. For context, in a review of narrative assessment tools used with three- to ten-year-olds, the number of stems ranged from four to fourteen (Bettmann & Lundahl, 2007). The smaller number of stems in this study may have decreased the measure's sensitivity to elicit and capture variance in children's data. Second, the coding scales for narrative coherence and behavioural dysregulation were adapted from existing scales, while a novel coding framework was developed to assess disruptive and prosocial story themes. While adaptations to coding approaches have been encouraged (Bretherton, Prentiss, et al., 1990), there is a risk of introducing unintended inconsistencies into the existing body of work (Woolgar, 1999). The

scales used in this study may benefit from further development or replication, particularly given the low internal consistency of the prosocial scale.

3.5.2.2.1.3 Resource-intensive data collection and analysis

The collection and analysis of story stem data was relatively time-consuming and costly. To administer the measure, researchers underwent half a day's training, completed three practice cases, and received ongoing supervision. Administration of the measure generally took 15 to 25 minutes. The research team were highly experienced, and qualities acquired through extensive fieldwork may have led to more efficient measure administration than would be seen in a less experienced team. To code the data, researchers received between two and four days of face-to-face training, with some additional time required to complete reliability sets. Coding took 40 to 60 minutes per case. The time and cost of data collection and analysis may be prohibitive for others looking to integrate such measures into their work. Indeed, a related limitation is that I both collected and coded data for some families meaning I was not blind to all participant characteristics. Due to the resource-intensive nature of data coding, it was not possible for graduate coders to complete all analysis themselves.

3.5.2.2.1.4 Furthering children's involvement

Given the rights-based approach of this research, the study could have been strengthened by two additions to the study procedure. First, child-friendly assent materials (e.g., illustrated information leaflets; O'Farrelly & Tatlow-Golden, in press) could have been used to explain the study's purpose and procedures to the child, and to formally ascertain the child's willingness to participate (Dockett & Perry, 2005; Ford, Sankey, & Crisp, 2007). Second, the children's involvement was restricted to the provision of data. They were uninvolved in the interpretation or analysis of their stories. Respondent validation techniques, including joint viewings of video clips (Flewitt, 2005) and group 'research conversations' (Truong & Mahon, 2012), have been employed in smaller, ethnographic research. Trialling the feasibility of such processes in larger, quantitative studies may enhance future findings.

3.5.2.2.1.5 Facilitating remote completion

The remote completion of the story stem measure was not supported in the current study. This led to the exclusion of 14 children (5%) from families where caregivers were unavailable for a home visit but provided data online or over the telephone. While the volume of lost data was relatively small, the ongoing COVID-19 pandemic means that exploring the use of digital technologies in the psychological assessment of young children is a priority.

Computerised versions of story stem completion tools have been tested in feasibility studies with children aged five to eight years (Minnis et al., 2006) but have not been trialled on a larger scale or with younger children.

3.5.2.2.2 Adult-reported measures

3.5.2.2.2.1 Multi-informant approach

This study benefitted from three measures of adult-reported child behaviour. This comprehensive assessment of behaviour was further strengthened by a multi-informant approach, with second informant data provided by 76% of children's teachers. The consistent concurrent association found across different measures and informants of child behaviour allows for a greater degree of confidence that the story stem measure offers a suitable tool to learn more about children's behaviour and perspectives.

3.5.2.2.2.2 Informant bias

Primary caregivers reported on their discipline practices, mood, anxiety, and child behaviour. Single informants reporting on multiple variables may introduce bias and inflate associations due to shared method variance (Collishaw, Goodman, Ford, Rabe-Hesketh, & Pickles, 2009). Acquiring reports from second caregivers could have partially mitigated this limitation and extended the study's multi-informant approach. However, participation of second caregivers in HSHS was reliant on the availability and willingness of another caregiver to not only participate in the study visits but also a fortnightly intervention of approximately three months duration. Because of this, participation of second caregivers in HSHS was relatively low, with only 17% of families comprised of two participating caregivers. Social desirability bias may have also affected parental reports particularly given that parental mental health and discipline style can be regarded as sensitive or contentious issues. Suggestions to reduce such effects often focus on the use of observational data but these can be costly, time-consuming, and impractical.

3.5.3 Implications and conclusions

This play-based story stem measure yielded useful data for the majority of participating children, proving to be developmentally suitable and highly feasible. Researchers were able to code the data reliably across key constructs of interest and modest but consistent associations were found with parent- and teacher-reported data. This study was exploratory, and the findings preliminary, but it adds to a growing body of evidence that places importance on children's capacity to provide valuable insight into their experiences. Further development

and utilisation of such methods with preschool-aged children would prove helpful in our understanding of their use. This research is particularly unique in that it studied a sample of children recruited based on early parent-reported externalising behaviour problems. Given the prevalence and far-reaching effects of enduring difficulties, finding unique ways to better understand and assess children's behavioural development in their earliest years may improve our ability to support families and prevent problems from becoming embedded.

From a methodological perspective, this measure was feasible to deliver as part of a wider assessment in which the child completed seven other research activities. Researchers and clinicians who may be discouraged by the time required for delivery of more traditional story stem batteries (typically one hour) may find the brevity of the current measure (15-20-minute delivery) more acceptable. Further study of this measure's psychometric properties, including ways to enhance internal consistency, and investigations of its test-retest reliability and construct validity would provide even greater understanding of its suitability and usefulness.

It must be reemphasised that, in acknowledging the feasibility and usefulness of this measure, the suggestion is not for its data to replace that provided by caregivers, teachers, and clinicians. Instead, the insight gained can be viewed as an important addition to adult-reported and observational measures. Including children's perspectives may provide a more holistic picture of their development across settings and contexts, possibly resulting in better estimates of outcomes.

CHAPTER 4: Exploring associations between early child behaviour, parent-child interactions, and children's later narrative representations

4.1 Chapter rationale

The previous study demonstrated that the administration of a dolls house play story stem assessment battery had utility when administered to a sample of three- and four-year-olds recruited based on early parental reports of behavioural difficulties. The measure was feasible to deliver in amongst a wider research assessment and showed correspondence with behaviours observed by children's parents and teachers.

The current study builds on the findings of the previous chapter by using a longitudinal design to explore associations between children's early experiences and their later narrative representations. First, I examine associations between children's early behaviour problems and four domains of their narrative representations measured two years later. Then, I consider associations between early parenting and children's later narratives representations. Parenting is considered from two perspectives. First, associations between early parental discipline practices and children's narratives are explored. Then, the effects of a video-feedback parenting intervention, delivered to caregivers when their children were toddlers, on children's narratives are examined. Integrating children's own perspectives into studies of these trajectories may enhance our understanding of child development, allowing us to make more targeted decisions in early prevention and intervention efforts.

4.2 Introduction

When exploring the developmental pathways of children's behavioural and emotional wellbeing we rely almost exclusively on information provided by adults (e.g., caregivers, educators, clinicians). Comparatively speaking, very few research studies include attempts to learn about how children themselves understand their early experiences. Such omissions of children's perspectives are especially true in studies of very young children. Understanding built on observations made by adult informants alone restricts our insight to only include child behaviours that are directly observable. Consequently, we forego the opportunity to gain an understanding of the internal systems and processes within the child that may be driving behaviour, as well as the environmental influences that shape these internal systems.

Utilising assessment tools that allow children's perspectives to be included in research studies is an important step in our understanding of developmental psychopathology. Efforts to gain such insight may be especially meaningful for children experiencing behavioural difficulties as it has been suggested that such difficulties may be related to children's internal ideas, beliefs, and feelings about socio-emotional problems (Zahn-Waxler et al., 1994). It is possible that individual differences in the presentation and persistence of behaviour problems could be reflected in aspects of these internal systems but adults around the child are limited in their capability to report on the intricacies of these inner processes. However, children's capacity to directly describe and discuss aspects of these systems themselves is constrained by their developing language, attention, and cognitive skills, and difficulties in conceptualising and discussing complex emotions. The use of representational methods that intend to tap information related to children's inner worlds offers one way to advance our understanding of this area. Utilising these approaches with children in their earliest years may be especially insightful for those designing and developing early prevention and intervention supports for families, something that has recently become a key focus of UK policy directives (Allen, 2011; Department of Health and Social Care and Department for Education, 2017).

Here, I reflect upon the theoretical underpinnings and methodologies of representational approaches, discuss ways in which previous research has identified links between measures of children's internal representations, their behaviour, and experiences of parenting, and outline the aims of the current study.

4.2.1 Theoretical basis to exploring children's internal representations

The precise ways in which 'inner worlds' are described and conceptualised differs across the field, with reference to internal representations, internal systems, internal working models, scripts, and schemata, among other terms. These phrases are sometimes used to describe the same idea while at other times to depict diverse concepts. Attachment is both a core theory and measure in the study of internal representations. Duschinsky et al. (2021) offered a comprehensive discussion of the convergence and divergence in the varied conceptualisations of attachment constructs across developmental psychology, social psychology, psychiatry, and psychotherapy. In their discussion, they compared the variance in the conceptualisations of attachment across psychology traditions to the issues of measurement invariance in statistics. They proposed that divergent forms of attachment discourse are "shaped by distinct social contexts, the priorities of these contexts, and their criteria for what counts as relevant and dependable knowledge and measurement, as well as criteria for who is considered a credible speaker" (Duschinsky et al., 2021, p. 2). In this section, I focus most closely on the conceptualisation of internal working models.

The term 'internal working models' was initially offered as a metaphorical concept, rather than a defined construct. Consequently, the idea has been used to explain a whole range of empirical associations, while "considerable uncertainty about its defining features, functioning, and measurement" remains (Thompson, 2018, p.332). Despite these enduring uncertainties, the concept of internal working models has acted as a foundation in a body of work conducted by both social and developmental psychologists (Pietromonaco & Barrett, 2000).

In developmental psychology, there is significant overlap between the study of attachment and the study of parenting. The meaning attributed to internal working models describes a set of expectations or predictions about the accessibility and responsiveness of key attachment figures (Society for Emotion and Attachment Studies, 2021; Thompson, 2007). While these expectations are thought to have some stability over time, they are not fixed but rather updated based on the processing of new experiences. These internal systems have been described as playing a role of calibration in the attachment behavioural system so it corresponds with the quality of care that a child is likely to receive based on prior experiences (Dykas & Cassidy, 2011). Under Bowlby's formulation, this set of expectations are then considered to extend more broadly to children's interpretation of their relational experiences, influencing how they approach, respond, and relate to people, relationships, and situations

(Ainsworth et al., 1978; Bowlby, 1969/1982; Bretherton, Ridgeway, et al., 1990; Erickson et al., 1985).

In social psychology research, internal working models have often been portrayed as elaborate, distinct, and independent symbolic representations of both the 'self' and 'others', contributing to the construction of the self (Fonagy, Gergely, & Target, 2007; Nelson, 1981). It has been suggested that individuals form cognitive representations that are functional and dynamic small-scale models of external realities, which act as 'reference points' when regulating behaviour and affect, processing information, and understanding others (Craik, 1943; Grossmann, 1999; Nelson, 1981; Huesmann, 1988). Thompson's ideas (Thompson, 2007, 2015, 2018) suggest that internal working models develop in accordance with social expectations, self-awareness, memory, and other social-cognitive skills, and are not only influenced by direct experiences but also through secondary representations formed out of conversational discourse with others.

As part of this, it has been proposed that internal working models play a role in social information processing. As children's cognitive functions advance, it is suggested that individuals with internal working models built on secure attachment histories are more likely to process information in a positive way. Conversely, those with models characterised by attachment insecurity are more likely to either exclude information from processing altogether, or do so with a negative bias (Dykas & Cassidy, 2011). This idea suggests that individuals' processing of information can shape their understanding of past, present, and future experiences, with internal working models acting as "representational bridges that mediate – through a variety of information processing mechanisms – longitudinal links between early experience and later adaptation" (Dykas & Cassidy, 2011, pg. 23).

There is significant overlap across these differing formulations meaning they are not necessarily incompatible with one another, despite the nuances in the ways in which the influence, malleability, and accessibility of internal working models are understood. A common thread running through them is the perception that internal working models "change developmentally with the child's conceptual advances and that social experience is formative to the development and potential revision of early [internal working models]" (Thompson, 2018, p.333). In this thesis, these concepts are discussed with an awareness that, while there is a rich theoretical basis for these ideas, more empirical work is needed to reach a fuller understanding of the exact processes underlying the formation of these systems, the specific

influences that shape them, and the extent to which they drive children's behaviour. In the current research, I mostly use the terms 'internal representations' and 'narrative representations' to describe concepts thought to be related to children's internal working models.

A pattern of empirical findings linking measures of children's attachment security and constructs thought to be related to their internal representations offer convincing support for the influence of attachment on children's inner worlds (see Yuval-Adler & Oppenheim, 2015 for a review). However, it is worth noting that this has not been a consistent finding across studies (e.g., Bretherton, Ridgeway, et al., 1990; Oppenheim, 1997). This suggests that children's internal representations are likely to also be influenced by, and representative of, other child characteristics, beyond attachment security, as well as wider family, peer, and sociocultural factors and processes. Evidence in support of the complex nature of these relationships comes from a recent set of meta-analyses that considered the formation and stability of attachment representations. The roles of the numerous moderating and mediating environmental, psychological, and biological mechanisms that underlie the relationships between attachment and child functioning were acknowledged across this research (e.g., Fearon et al., 2010; Groh et al., 2014; Groh et al., 2012; Verhage et al., 2016).

In sum, despite many decades of theoretical and empirical research, the precise mechanisms underlying the conceptualisation, definition, and formation of children's inner worlds are still under investigation and open to debate today. Many studies have focused exclusively on children's attachment representations as a measure of internal systems, while others have considered broader social, emotional, or behavioural components of children's inner worlds. Through this research, I draw on thinking related to both social and developmental psychology traditions to develop research questions and interpret findings. Next, I consider the methods and approaches that have been used to learn more about these inner worlds.

4.2.2 Methods and approaches to accessing internal representations

There is variance in the methodological approaches undertaken to elicit and measure aspects related to internal representations.

One of the most common constructs included in studies of internal representations is attachment. Approaches to measure attachment security vary. For instance, interviews of attachment representations have been developed for use with adults (e.g., the Adult

Attachment Interview; AAI, George, Kaplan, & Main, 1996; Main, Kaplan, & Cassidy, 1985), and children in middle childhood and adolescence (e.g., the Friends and Family Interview, FFI, Steele & Steele, 2005; the Child Attachment Interview, Shmueli-Goetz, 2014; Shmueli-Goetz, Target, Fonagy, & Datta, 2008). Such methods are considered to elucidate states of mind with respect to attachment (Hesse, 2008) and usually require individuals to describe their relationship with a significant other (i.e., parents, children, siblings, teachers, best friends; George et al., 1996; Kriss, Steele, & Steele, 2012) using detailed examples. Typically, in these methods the overall narrative coherence of speech (e.g., succinctness, relevance, clarity, contradictoriness) has been used as a marker of security in attachment representation, as opposed to the thematic content of the recall. Participation in such procedures is reliant on metacognitive and memory skills to enable the recollection of past interactions that are representative of an individual's higher-order perceptions of themselves and others (Steele & Steele, 2005). Such a skillset is primitive in very young children and thus using strictly verbal interview methodologies could elicit unreliable data and significantly reduce the extent to which young participants can offer their perspectives. Observational methods such as the Strange Situation Procedure have been used to assess the attachment representations of infants and preschoolers (Ainsworth et al., 1978; Cassidy & Marvin, 1992). These structured observations are designed to activate the child's attachment system, allowing for the observation of the child's attachment behaviours during separations and reunions with their caregiver, and a classification of attachment security to be assigned (Society for Emotion and Attachment Studies, 2021).

The utilisation of these interview and observational methods in previous research has provided helpful insight into children's representations of their key attachment figures, allowing us to better understand the place of attachment in children's developmental trajectories. However, the classification of attachment offers only one way to better understand individuals' internal systems. Attachment classifications are limited in their ability to attain more detailed information related to how children understand and perceive not only their relationship with their caregiver(s) but also the situations around them and their own behaviour. Measures aiming to access broader components of children's internal representations and emotional processes have also been used to better understand the socioemotional development of children (e.g., Bretherton & Munholland, 2008; Robinson, 2007). One approach that considers these constructs in more detail uses children's narratives

as a measure of their representational worlds. Relatively speaking though, this is an under-utilised approach in studies of child behaviour. Their use is considered in more detail next.

4.2.2.1 Using narrative-based representational approaches to measure internal representations

Semi-projective measures have been used with young children to attain insight into their internal perspectives. These techniques are built on the assumption that, when creating imaginary responses, children draw on real-life experiences in their responses to semi-projective activities (Bretherton & Munholland, 2008).

Doll play story stem methodologies have been used to explore aspects of children's early experiences and representations of their internal and external worlds, using play as a medium of communication that is familiar and developmentally sensitive (Woolgar, 1999; Yuval-Adler & Oppenheim, 2015). Such narrative methods are considered to minimise issues related to a child's desire to give socially acceptable answers by engaging the child to such an extent that they activate script-like representations that underlie components of typical behaviours (Murray et al., 1999). The process of presenting the child with a 'dilemma' story beginning and inviting them to show and tell what happens next is believed to act as a catalyst for the child to draw upon the knowledge and understanding of similar situations provided by their internal representational models (Solomon & George, 2008).

Story stem batteries have been administered and coded in a variety of ways but broadly-speaking codes fall into three domains: the organisation and structure of the story, the content of the story, and the child's behaviour during the task (Page, 2001). Analysis of the child's narrative response is considered to reveal important dimensions of their internal systems, with clear and complete narratives reliant on an ability to regulate emotions in order to explain thoughts in a coherent way (Robinson, Hérot, Haynes, & Mantz-Simmons, 2000). Despite their complex nature, these skills begin to develop at two years old and by three years of age, when given a structured framework, props, and prompts, most children can produce narratives related to a dilemma (Bretherton, Ridgeway, et al., 1990). Several studies have supported the validity of such approaches (Bretherton & Oppenheim, 2003; Green et al., 2000; Page, 2001). The findings of the previous chapter – where 94% of three- and four-year-olds completed the story stem measure, and modest but consistent associations were found between children's narrative responses and corresponding parent- and teacher-reported data – also contribute to our understanding of the suitability of such techniques.

4.2.3 Links between external and internal worlds

Children's responses to story stem methodologies have been used as a measure of their internal worlds in a number of investigations of child development. These studies have revealed numerous links between aspects of children's narratives and factors related to their social, emotional, and behavioural functioning and wider family and social environments (see Yuval-Adler & Oppenheim, 2015 for a review). Here I consider just two of these factors: child behaviour and parental discipline.

4.2.3.1 Behaviour and children's narrative representations

Components of children's story stem responses have been linked to wider aspects of their functioning, including associations with emotional difficulties, peer problems, school adjustment, and their self-esteem (Yuval-Adler & Oppenheim, 2015). Of particular interest to the current set of studies are relations between children's narratives and their behavioural functioning, due to the pervasive effects of enduring behaviour problems discussed previously in Chapters 1 and 3. In the literature, parent-reported measures of behavioural difficulties have been linked to several aspects of children's narratives including disruptive, aggressive, antisocial, and dysregulated thematic content (e.g., Futh, O'Connor, Matias, Green, & Scott, 2008; Moss et al., 2009; Stacks, 2007; von Klitzing et al., 2000; Woolgar et al., 2001). However, children's story characteristics and the patterning of associations with observable behavioural difficulties have mostly been explored in non-clinical, low-risk samples. Children with behaviour problems are left understudied (Wan & Green, 2010). Those studies that have included children experiencing behavioural difficulties have generally been restricted by small sample sizes (e.g., Hill et al., 2007: $N = 41$; Hubbs-Tait et al., 1996: $N = 27$; Wan & Green, 2010: $N = 77$; Zahn-Waxler et al., 1994: $N = 89$). Further to this, the majority of these studies have been concurrent in nature, with only a minority adopting a longitudinal design (Yuval-Adler & Oppenheim, 2015). In a sample of six-year-olds, parent-reported attentional difficulties were associated with children's narrative representations of their caregivers measured one year earlier (Davies, Woitach, Winter, & Cummings, 2008). In one of the only studies to include preschool-aged children, Warren, Oppenheim, and Emde (1996; $N = 51$) reported cross-lag associations between children's narrative themes and behavioural difficulties for four- and five-year-olds but not for three-year-olds. To my knowledge, there are no studies that adopt a longitudinal design to explore associations between child behaviour problems exhibited in toddlerhood and preschool-aged children's later narrative representations in a sample considered to be clinically at-risk.

4.2.3.2 Parental discipline and children's narrative representations

The concept of internal working models is underpinned by the belief that children's earliest interactions with those around them shape the organisation and content of their internal systems. Given this, it is unsurprising that characteristics of parenting are a core focus in the story stem field. Significant emphasis has been placed on links between children's narratives and the quality of attachment in the parent-child relationship, although such studies with preschool-aged children are in the minority (e.g. Miljkovitch, Pierrehumbert, & Halfon, 2007; Oppenheim, 1997). To a lesser extent, other elements of parenting have also been considered including communication style (e.g., Reese et al., 2020), maltreatment (e.g., Toth et al., 1997), parental supportiveness (e.g., Berzenski & Yates, 2017), and parents' own exposure to violence and trauma (e.g., Schechter et al., 2007). Here I focus on a single component of parenting: discipline style.

Discipline is an inescapable and necessary feature of parenting. Discipline in and of itself is not 'bad' but the style and approach of caregivers engaging in discipline-related interactions can be significant for children's outcomes. Effective and productive approaches to discipline have been associated with lower behavioural difficulties (Choe et al., 2013), and a greater capacity for self-regulation (Karreman et al., 2006) and prosociality (Knafo & Plomin, 2006). Effective discipline strategies have been said to place caregivers in the role of external regulators, offering support and guidance to children as they learn to regulate their emotions and inhibit disruptive behavioural patterns (Choe et al., 2013; Kochanska & Aksan, 2006). Conversely, the detrimental effects of negative, inconsistent, and harsh parental discipline strategies on children's behavioural and emotional outcomes have been established across childhood, including in the earliest years (Chang et al., 2003; O'Leary & Vidair, 2005; Parent et al., 2016; Weiss, Dodge, Bates, & Pettit, 1992).

Most of this research has focused on associations with children's observable, external behaviours with less focus on the assessment of links to children's internal representations. Studies that have used doll play methodologies in the context of parental discipline have predominantly coded children's own representations of discipline in their stories (e.g., Belden, Sullivan, & Luby, 2007; von Klitzing et al., 2007). However, an alternative (less-studied) approach is to directly measure children's exposure to parental discipline style and explore associations to other aspects of children's narratives. For instance, in a sample of low-risk six-year-olds, Laible et al. (2004; $N = 74$) reported concurrent associations between harsh parenting and children's use of prosocial and aggressive story themes but not their narrative

coherence. Mackenbach et al. (2014) reported that parental discipline at three years old was associated with children's own reports of their behaviour using a puppet interview at six years old. The available research predominantly focuses on the representations of older (5+ years), non-clinical, low-risk children, measured in a concurrent manner. We are lacking in an understanding of the early influences of parental discipline on the ways in which children themselves understand and represent their worlds. Such understanding may not only help us better understand the developmental mechanisms of children's psychological functioning in their earliest years but also enhance our recognition of suitable targets of early prevention and intervention efforts to promote child and family functioning.

4.2.3.3 Parenting interventions and children's narrative representations

One especially unique and potentially insightful way to extend upon our understanding of the influences of early parenting on children's representations is through the examination of interventions designed to promote positive parenting practices. As the quality and content of parent-child interactions are considered to feed into children's internal systems, it is reasonable to assume that programmes seeking to enhance parenting practices may also affect aspects of children's narratives, as a reflection of their internal representations. There is a very small body of empirical research indicating that programmes that aim to enhance features of parenting may have knock-on effects on the ways in which children understand and conceptualise their worlds. Olds et al. (2004) explored the narratives of six-year-olds in a follow-up study of the Family Nurse Partnership programme (Olds, 2006). In families categorised as 'low-resourced', children whose parents received the programme told less aggressive (d = -0.25) and less incoherent (d = -0.34) stories in response to the narrative prompts. Reese et al. (2020; N = 76) examined the narratives of 11-year-olds following a parenting programme aiming to enhance the quality of parent-child conversations. Children whose parents had participated in the intervention exhibited more coherent and elaborative narratives (ŋp2 .10 - .12; medium-sized effects). Toth, Maughan, Manly, Spagnola, and Cicchetti (2002) measured children's narratives at five years old, following randomisation to preschooler-parent psychotherapy (n = 23), psychoeducational home visitation (n = 34), or community standard (n = 30). Following the intervention, children in the preschooler-parent psychotherapy group exhibited a decline in maladaptive maternal representations and negative self-representations in their stories.

Yuval-Adler and Oppenheim (2015) note that the use of story stem techniques in studies supported by the methodological strengths of randomised controlled trials "is of great

importance in establishing direction of effects, and more studies using such designs are clearly needed" (p.369). Using an experimental design to explore the narratives of very young children following the delivery of a parenting intervention, in a window of time considered to be especially sensitive to the development of children's internal representations, could not only provide us with insight around influential parenting practices for children's representations but also the power of early intervention.

4.2.4 Conclusion and current study's research aims

The current study aims to extend the findings of the previous chapter by furthering our understanding around how early experiences may be associated with components of children's internal representations. There is evidence to suggest that aspects of children's narrative representations are influenced by their behavioural functioning and experiences of parenting. However, this evidence base is largely restricted to concurrent studies of low-risk, school-aged populations, with a sole focus on representations as a measure of attachment security. Using a more multifaceted measure of children's internal representations, exploring these relationships over time using a longitudinal design with a group of children aged one to five years could allow us to better understand the early predictors of the ways in which children come to understand and perceive their worlds.

Focusing on populations of children considered to be at an increased likelihood of developing behavioural difficulties may be especially important. Early behaviour problems may disrupt the coherent organisation of children's internal systems, and result in representational models that are characterised by their disruptive behavioural patterns. In turn, challenging child behaviour may affect the ways in which parents respond to their child (a relationship that is likely bidirectional; Danforth, Anderson, Barkley, & Stokes, 1991) and, of specific interest for the current study, may elicit dysfunctional discipline strategies. Consequently, these parental exposures may influence the content and organisation of children's internal working models and in turn their capacity to respond to 'real life' situations. Parenting interventions that promote positive aspects of parenting may improve the parent-child relationship, in turn enhancing aspects of children's internal systems and responses to their worlds. Figure 7 visually depicts these proposed relationships, although in reality such associations are most certainly more complex and multi-faceted. Central to the current study is the assumption that story stem methodologies can elicit aspects of the ways in which children internalise and process information to be drawn upon in their responses to external situations.

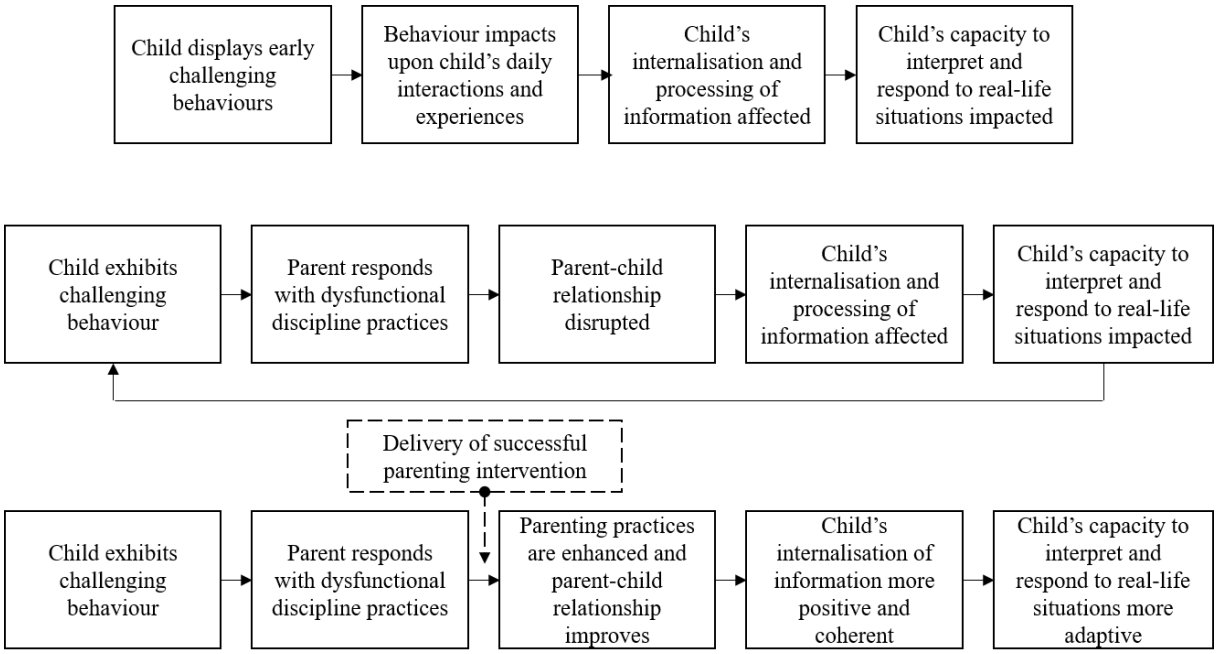


Figure 7. A visual depiction of the ways in which children's behaviour, parenting exposures, and internal representations may be related to one another.

To investigate these possible relationships further, the current research is driven by two key aims:

1. To investigate whether early child behaviour problems are associated with components of children's internal representations, measured two years later using a story stem battery.
2. To investigate whether early parenting is associated with children's later internal representations, measured using the story stem outcomes. Parenting will be considered in two ways:
    a. Associations between early overreactive and lax parental discipline and constructs derived from children's narratives will be explored.
    b. Effects of a video-feedback positive parenting programme on constructs derived from children's narratives will be tested.

Using measures that aim to access the internal worlds of very young children may allow us to understand if and how children's early behavioural difficulties are reflected in their internal representations. Exploring parenting as a key component to this may allow us to better understand the influential role of parents on children's functioning and the ways in which families can be supported in responding to their children.

4.3 Method

4.3.1 Study design

Data for this study were collected in the broader context of the Healthy Start, Happy Start trial (HSHS; published protocol: Ramchandani et al., 2017). The current study used a longitudinal design to explore associations between early child behaviour and parental discipline practices, measured when children were one and two years of age, and children's representations and behaviour during a doll's house play story stem assessment two years later. The study also utilised the experimental design of the HSHS study by investigating whether a video-feedback parenting intervention aiming to promote positive parenting practices was associated with children's outcomes measured through the story stem assessment.

I was involved in all aspects of data collection in the HSHS study, completing 38% (*n* = 327) of research assessments undertaken with families. I also co-led on the substantial ethical amendment required for the current study and the piloting of included measures and led on the training and coding required for data analysis in this research.

4.3.2 Participants

Participants were children and primary caregivers who completed both their baseline and 24-month follow-up assessment in the HSHS trial. Detailed information about participant recruitment is provided in Chapter 2, section 2.3.4. In brief, in six NHS sites across the Southeast and East of the UK, caregivers of children aged 12-36 months were asked to complete a screening questionnaire (SDQ; Goodman, 1997) for early behaviour problems. Families whose children scored in the top 20% of population norms for externalising difficulties were invited to take part in the main trial. Following enrolment into the study, families were randomised into either the treatment (offered six sessions of VIPP-SD plus access to usual care) or control (access to usual care offered by health, social, and community services alone) arm of the trial. Families were followed up twice after their baseline visit, at five months and 24-months post-randomisation.

Data for 300 participating families were collected at baseline, and 282 families at the 24-month follow-up assessment. Of these, 256 children fully completed the story stem assessment battery (the primary focus of the current study). Reasons for non-completion of the story stem measure were as follows: assessment visit conducted over telephone instead of

face-to-face (n = 14); child did not assent to measure completion (n = 11); and child only partially completed the measure (n = 5). There was minimal other missing data in this sub-sample, with only one caregiver not completing a questionnaire measure of child behaviour and another not completing a measure of parental depressive symptoms at baseline.

4.3.3 Measures

For brevity, measure descriptions in this section are written concisely as more detailed information is provided in Chapter 2, Section 2.3.8. Figure 8 provides a visual depiction of the timing of data collection for this study.

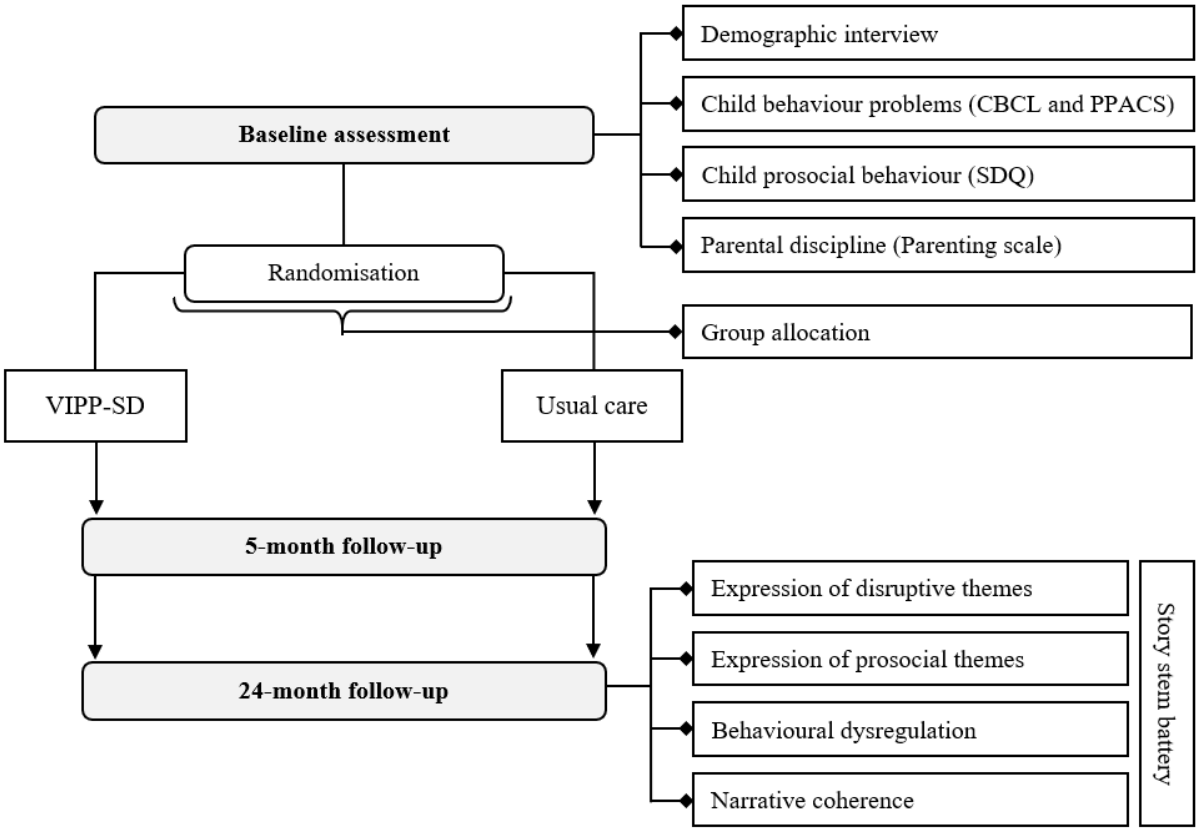


Figure 8. Timing of data collection for the examination of longitudinal associations between early child behaviour and parenting environment and later narrative representations.

4.3.3.1 Demographic measure

Demographic information for participating caregivers and children was collected through a researcher-led interview with the child's primary caregiver. Demographic variables included caregiver age, sex, ethnicity, educational attainment, employment status, and relationship status, and child age, sex, and ethnicity.

4.3.3.2 Child informant measure: Doll's house play story stem assessment battery

The doll's house play story stem assessment battery was administered to capture constructs related to children's internal representations and perspectives of their behaviour. Section 3.3.3.1 of Chapter 3 provides extensive detail about the adaptation and piloting of this measure, as well as the development of the coding scheme used to analyse children's responses. For conciseness, an abridged overview of the development and administration of this measure is provided below.

4.3.3.2.1 Story stem measure development

The story stem measure administered to the HSHS children at the 24-month follow-up was adapted from and informed by two existing approaches that made use of doll play techniques: the MacArthur Story Stem Battery (MSSB; Bretherton & Oppenheim, 2003; Oppenheim, Emde, & Warren, 1997) and the Dolls House Play approach (Murray et al., 1999; Ramchandani et al., 2011; Woolgar & Murray, 2010). Through careful piloting, three story stems from these measures were adapted for use in the HSHS study. Piloting of the measure took place in two phases. Phase one involved administering the measure to 15 three- and four-year-olds in two nursery school settings. Following adaptations from the first stage of piloting, the second phase involved administering the story stem battery to five children in their homes, as part of a wider piloting of the complete two-year follow-up assessment.

4.3.3.2.2 Story stem procedure

Three researchers (including myself) conducted the story stem assessment with participating children. This measure was administered as part of a wider research assessment lasting approximately two and a half hours in the family home. For the story stem assessment, the researcher sat with the child at a table or on the floor where a table was not available. The researcher explained to the child that they were going to tell three stories together, and that the researcher would tell the beginning and the child could tell the ending. If the child was happy to take part in the activity, they were asked to pick dolls that they thought were most like them and the family members that lived with them. Following a warm-up story (a party) that introduced the child to the core concepts of the activity, the child was invited to play through the three scenarios (spilt juice, hurt knee, three's a crowd). In each of these scenarios, the doll representing the child was placed in a situation of distress or conflict. The child was then encouraged to finish the story through play and narration with the prompt "show me and tell

me what happens now" from the researcher. See Appendix E for the measure's administration manual.

4.3.3.2.3 Coding of the story stem data

I coded the story stem data with the support of three graduate students. All coders reached moderate to excellent levels of inter-rater reliability (ICC: .68-.98) based on guidance thresholds of reliability (Koo & Li, 2016). The coding scales applied to the story stem videos were informed by coding approaches undertaken in other studies that have used children as informants. A full overview of the development of the coding scheme is provided in Chapter 3, Section 3.3.3.1.4. The scoring used in the current study was divided into four core sections:

1. *Disruptive content themes:* Expression of aggressive, destructive and ostracising themes in the story.
2. *Prosocial content themes:* Expression of helpful, kind, caring themes in the story.
3. *Narrative coherence:* How organised and coherent the child's narrative was.
4. *Observed behavioural dysregulation:* Ability to regulate behaviour during measure administration (e.g., staying in seat, maintaining concentration).

4.3.3.3 Parent-rated measures of behaviour

In the previous study, where the analysis of the story stem data was particularly exploratory in nature, child behaviour was assessed using three different parent-reported measures, and both the total and subscale scores of each of these measures were examined. However, to focus the current study, and to limit the number of statistical tests, two measures of child behaviour problems and one measure of prosocial behaviour were selected. The reason two measures of behaviour problems were used in this analysis was to provide a more in-depth picture of children's behaviour at one and two years of age, particularly given the inherent challenges in assessing behaviour problems in toddlers and the novel nature of the current research questions. Each of these three measures are now briefly described.

4.3.3.3.1 Child behaviour problems: CBCL

The Child Behaviour Checklist (CBCL/1.5-5; Achenbach & Rescorla, 2000, 2001) is a measure of children's behavioural, emotional, and social problems. The 100-item questionnaire asks parents or caregivers to rate how true they feel each behaviour is of their child in the last two months on a three-point scale (0 = not true, 1 = somewhat or sometimes true, or 2 = very true or often true). The measure provides a total score, an externalising score,

and an internalising score. The CBCL is a widely used and well-validated measure (Achenbach & Rescorla, 2000; Rescorla et al., 2011). A key strength of this measure for the current study is that it has been validated for use with children as young as one year old (van Zeijl et al., 2006), which is important given that children were 12-36 months old at the time of the baseline assessment.

4.3.3.3.2 Child behaviour problems: PPACS

The Preschool Parental Account of Childhood Symptoms (PPACS; Sonuga-Barke et al., 1994; Taylor et al., 1991) is a semi-structured researcher-led interview of children's attentional and conduct difficulties. In this study, the measure was conducted with the child's primary caregiver and generally took between 45 minutes and 1 hour to complete. During the interview, the caregiver was asked to provide detailed descriptions of their child's behaviour in the last week. They were then asked to report on how representative these examples were of the child's behaviour over the last four months. The researcher used operationalised criteria for each of the behaviours, as well as their clinical judgement and training, to rate the child's attention and conduct difficulties on scales of severity and frequency. Across the study, the interviews were conducted by five researchers (including myself). One-way, single-measurement, absolute agreement intra-class correlation coefficients indicated excellent levels of inter-rater reliability between assessors at all timepoints (ICC: .92-.98). A strength of this measure is that interview methodologies have been regarded as the 'gold-standard' in developmental psychopathology (National Institute for Health and Care Excellence, 2013). However, compared to the CBCL, the PPACS has been less widely used with children aged under three years. Thus, the current study is strengthened by having access to both a clinical interview and a well-validated questionnaire assessment of child behaviour in one- and two-year-olds.

4.3.3.3.3 Child prosocial behaviours: SDQ

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a 25-item questionnaire designed to measure children's behaviour across five subscales (conduct problems, emotional symptoms, hyperactivity/inattention, peer relationship problems, and prosocial behaviour). Respondents rate each item on a three-point scale (0 = not true, 1 = somewhat true, or 2 = certainly true) based on how true the behaviour is of the child in the last six months. In this study, an adapted version of the SDQ suitable for use with preschool populations was used. Only the prosocial subscale was included in this analysis. Studies of

this measure's psychometric properties when administered to toddler- and preschool-aged populations have reported adequate model fit and acceptable internal consistency (α range = 0.45–0.82) of the scales (Croft et al., 2015; Doi, Ishihara, & Uchiyama, 2014; Ezpeleta et al., 2013; Klein et al., 2013; Theunissen, Vogels, de Wolff, & Reijneveld, 2013). Concentrating on the prosocial subscale, it has been found to have good internal consistency (father, mother, and teacher-reported αs = 0.82-0.92; 3-5 years; Dahlberg, Ghaderi, Sarkadi, & Salari, 2019), test-retest reliability (r = 0.87; 1-3 years) and concurrent validity (r = 0.58; 1-3 years) (Gustafsson, Gustafsson, & Proczkowska-Björklund, 2016) in studies of one- to five-year-olds.

4.3.3.4 Parent-rated measures of parenting and mood

4.3.3.4.1 Parental discipline practices: Parenting Scale

The Parenting Scale (Arnold et al., 1993) was used to measure caregivers' approach to discipline with their child. Participating caregivers were asked to respond to 30 items by rating their parenting behaviour on a seven-point Likert scale. For each item, a situation related to children's misbehaviour is presented alongside an 'effective' and 'ineffective' approach to addressing the situation at either end of the scale. Three subscale scores for laxness (permissive discipline), overreactivity (harsh discipline), and verbosity (overly wordy responses to misbehaviour), as well as a total score, are produced. Studies that have attempted to replicate the factor structure of the scale originally established in the questionnaire's development have consistently failed to replicate the verbosity subscale (Harvey et al., 2001; Reitman et al., 2001; Steele et al., 2005). Due to these issues, in the current study the laxness and overreactive scales were analysed separately and the verbosity scale was not included in analysis.

4.3.3.4.2 Parental low mood

Parent mood was assessed through the Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001). This self-reported, nine-item questionnaire measures the symptomology and severity of depression. Respondents are asked to report on how often in the last fortnight they have experienced the listed symptoms on a scale of 0 (not at all) to 3 (nearly every day). A total score is obtained by summing all items of the questionnaire.

4.3.4 Study procedure

Study assessments were carried out in families' homes or, in a minority of cases, a private room in a nearby community centre. Participating caregivers provided written informed consent for themselves and their child at each phase of the study. Following consent, the primary caregiver completed the clinical interview (PPACS) with the researcher. Several play-based interactions between the child and caregiver were then video recorded. Following this, caregivers completed a battery of questionnaires (including the CBCL, SDQ, Parenting Scale, and PHQ9) while the researcher administered direct assessments of the child (including the story stem assessment battery).

4.3.5 Intervention: VIPP-SD

Families were randomly allocated to one of two groups following their baseline assessment. The randomisation procedures and conditions of each group are discussed here.

4.3.5.1 Randomisation

Participants were randomly allocated (1:1) to either the intervention (offered six sessions of VIPP-SD plus usual care) or control arm (usual care alone, i.e. access to general health and community supports). The randomisation list was generated by an independent statistician and used block sizes of 2, 4, and 6 varying at random. Randomisation was stratified by recruitment site and number of participating caregivers (one versus two). Participating families were allocated to the next available treatment code. Access to the allocation sequence was restricted and not accessible to any members of the research team.

4.3.5.2 Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD)

Families randomised to the treatment arm of the trial were offered six home-based sessions of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer et al., 2008b) parenting programme. The intervention is a brief, manualised, home-based intervention that aims to promote parental sensitivity and a sensitive discipline style. Intervention visits were scheduled on an approximately fortnightly basis and were generally one to two hours in duration. Therapists delivering the programme were predominantly from health visiting, nursing, and psychology backgrounds. In the visits, videos of parent-child dyads in everyday situations were used by the therapist to offer the caregiver feedback on aspects of their child's development and the parent-child relationship.

Feedback was guided by a 'script' prepared in advance by the therapist, which was informed by the sensitivity and discipline themes outlined in the intervention manual.

4.3.5.3 Usual care

Participants in both the VIPP-SD and usual care arms continued to access usual care. In most cases, this care is minimal as there are no standard care pathways through the NHS for early onset behaviour problems. General support available to families was mostly through health visiting and general practitioner services.

4.3.5.4 Blinding

Research assistants were 'blind' throughout data collection and analysis. In the rare instances where unblinding did occur (2% and 4% of assessments at each follow-up point), recordings of researcher-rated measures were double scored by a second, blinded assessor.

4.3.6 Statistical analysis

Analysis for this study was undertaken in four steps. First, the sample characteristics and descriptive statistics of key study variables were produced. Boxplots, histograms, normality plots (P-P and Q-Q plots), and tests for skew and kurtosis were used to examine the distribution of study variables and to detect outliers.

Following this, bivariate (Pearson) correlation coefficients were run to explore the stability of child behaviour and parenting over two time points (baseline and 24-month follow-up), and to explore associations between baseline measures of child behaviour and parental discipline, and children's outcomes on the story stem assessment two years later.

Next, assumptions of multivariate regression were tested. The G*Power tool (Faul et al., 2009) was used for an a priori power analysis calculation. For linear multiple regression, to detect a medium effect size ($f^2 = 0.15$), a sample size of 114 was required ($\alpha = 0.05$; at 80% power, with up to nine predictors). Based on this analysis, the sample size of this study ($N = 256$) appeared to be sufficient to detect effects of a medium magnitude. Multiple linear regressions were conducted to explore the associations between early child behaviour and parental discipline style and outcomes measured through the story stem assessment. Story stem scales (disruptive themes, prosocial themes, narrative coherence, and self-regulation) measured at 24-months post-randomisation were the dependent variables and were each examined in separate regression models. Baseline predictor variables of interest were early behaviour problems (measure 1: CBCL total, measure 2: PPACS total), prosocial behaviours

(SDQ prosocial subscale), and parental discipline (Parenting Scale laxness and overreactivity subscales). This analysis allowed for the control of key variables. In the previous chapter, child gender and age, and parental educational attainment and depressive symptoms, were found to be significantly associated with aspects of children's story stem outcomes and so were controlled for. Controlling for treatment allocation (VIPP-SD or usual care) also allowed for the control of any effects of the VIPP-SD intervention on children's outcomes.

Finally, independent samples *t*-tests were used to examine differences in outcomes from the story stem assessment by treatment allocation, to explore whether there were any intervention effects on children's narrative representations. All participants were analysed according to the treatment arm they were allocated to, irrespective of whether they received any or all of the prescribed treatment dosage (six VIPP-SD sessions).

Approach to missing data can be found in Chapter 2, section 2.3.14. Analysis was completed using SPSS version 25.

4.4 Results

This study investigated associations between early child behaviour and parenting (measured at 12 to 36 months old) and features of children's narrative representations during a story stem assessment (measured at three to five years old). Parenting was considered in two ways in this study. First, using an observational, longitudinal design, the predictive relationship between early parental discipline style and children's later narrative representations was explored. Second, an experimental design was utilised to test the effects of a video-feedback parenting intervention, aiming to promote parental sensitivity and a sensitive discipline style, on features of children's narrative representations.

The results are presented as follows: First, for the preliminary analysis, the sociodemographic characteristics and data completeness of the sample are discussed and the associations between predictor and outcome variables across the two time points are explored. Second, for the main analysis, multivariate assumptions are tested before multiple regressions, run to control for covariates in the exploration of the associations between early child behaviour and early parental discipline and late story stem outcomes, are presented. Finally, group differences on children's narrative representations according to group allocation (VIPP-SD vs. usual care) are assessed.

4.4.1 Preliminary analysis

4.4.1.1 Sample characteristics

The sociodemographic characteristics of this sample are provided in Table 21 and the flow of participants through the two years of the study is outlined in Figure 9. The current sample included only children who fully completed the story stem assessment (n = 256), excluding the children who only partially completed the measure (n = 5). There was a small amount of missing data at the baseline assessment, with one caregiver not completing the questionnaire for depressive symptoms (PHQ-9) and one not completing the measure of prosocial behaviours (SDQ subscale). This means there is some small variance in sample size between different analyses presented in this chapter (ns 254-256).

At baseline, children were aged between 12 and 36 months (M = 23.0 months; SD = 6.7 months). At the two-year follow-up, children were aged between 36 and 65 months (M = 48.0 months; SD = 7.1 months). Slightly more male than female children took part (53% male), and most children were either white (65%) or of Mixed heritage (21%). Most participating caregivers were female (96%), the child's biological or adoptive mother (96%), held a

graduate-level qualification (67%), and were either married, in a civil partnership or living with a partner (87%). Tables 51 and 52 in Appendix H detail the sociodemographic profiles of families and descriptive statistics of study variables in this subsample compared to the full HSHS sample.

Table 21

Sociodemographic Characteristics of the Full Sample and Split by Group Allocation

Characteristic	Trial arm		
	VIPP-SD (*n* = 128)	UC (*n* = 128)	All (*N* = 256)
Child characteristic			
Sex (female), *n* (%)	64 (50)	56 (44)	120 (47)
Age at baseline (months), mean (SD)	22.8 (7.0)	23.6 (6.5)	23.2 (6.7)
Age at 2YFU (months), mean (SD)	47.5 (7.3)	48.1 (6.8)	47.8 (7.1)
Ethnicity, *n* (%)			
Asian	8 (6)	6 (5)	14 (5)
Black	2 (2)	13 (10)	15 (6)
Mixed ethnicity	32 (25)	22 (17)	54 (21)
White	83 (65)	83 (65)	166 (65)
Other	3 (2)	4 (3)	7 (3)
Primary caregiver characteristic			
Sex (female), *n* (%)	122 (95)	123 (96)	245 (96)
Age at baseline (years), mean (SD)	34.6 (5.1)	34.7 (5.7)	34.7 (5.4)
Parental status, *n* (%)			
Biological or adoptive mother	123 (96)	123 (96)	246 (96)
Biological or adoptive father	5 (4)	5 (4)	10 (4)
Ethnicity, *n* (%)			
Asian	13 (10)	13 (10)	26 (10)
Black	3 (2)	13 (10)	16 (6)
Mixed ethnicity	9 (7)	9 (7)	18 (7)
White	95 (74)	92 (72)	187 (73)
Other	8 (7)	1 (1)	9 (4)
Relationship status, *n* (%)			

Characteristic	Trial arm		All
	VIPP-SD (n = 128)	UC (n = 128)	(N = 256)
Married/civil partnership/cohabiting	111 (87)	111 (87)	222 (87)
Single	9 (8)	12 (9)	21 (8)
In relationship but not cohabiting	7 (6)	1 (1)	8 (3)
Divorced/widowed/legally separated	1 (1)	4 (3)	5 (2)
Employment status, *n* (%)			
Employed	74 (58)	66 (51)	140 (55)
Looking after home and children	46 (36)	45 (35)	91 (36)
Paid parental leave	6 (5)	10 (8)	16 (6)
Full-time student	2 (1)	7 (6)	9 (3)
Highest qualification, *n* (%)			
Graduate	84 (66)	88 (69)	172 (67)
A-level/NVQ/BTEC	31 (24)	30 (23)	61 (24)
GCSE or lower	13 (10)	10 (8)	23 (9)

Note. VIPP-SD = Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline; UC = usual care.

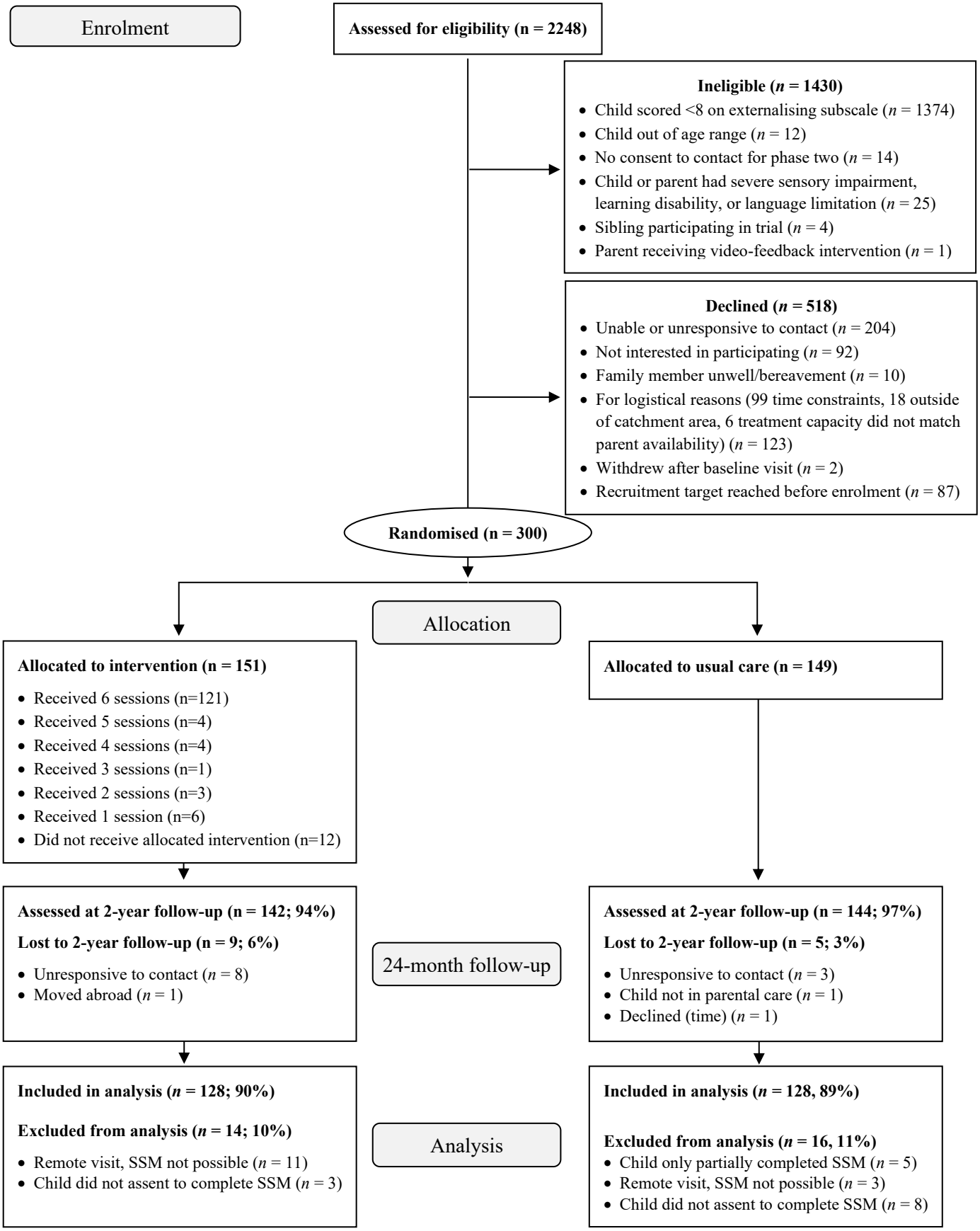


Figure 9. The CONSORT flow diagram for participant data included in the current analyses.

Note. SSM = story stem measure.

4.4.1.2 Descriptive statistics of study variables

Descriptive statistics of study variables are presented in Table 22. Measures of child behaviour (PPACS, CBCL, and SDQ), parental discipline (Parenting Scale) and parental mood (PHQ-9) were collected at the baseline assessment. Variables obtained through the story stem measure (disruptive and prosocial story themes, narrative coherence, and behavioural dysregulation) were collected at the 24-month follow-up assessment.

Table 22

Descriptive Statistics for Study Variables Included in the Current Analyses

Variable	*N*	M	SD	Min	Max	Skewness	Kurtosis
Baseline assessment							
CBCL – total	256	40.55	20.64	3.00	108.84	0.91	0.96
PPACS – total	256	32.59	9.39	8.89	65.00	0.37	0.10
SDQ – prosocial subscale	255	5.59	2.23	0.00	10.00	-0.26	-0.27
Parenting scale - overreactivity	256	3.89	0.46	2.60	5.20	-0.13	-0.08
Parenting scale - laxness	256	3.69	0.43	1.82	4.91	-0.18	1.19
PHQ-9 - total	255	3.96	3.66	0.00	18.00	1.17	1.08
24-month assessment							
Disruptive story themes	256	5.00	4.10	0.00	19.00	1.06	0.66
Prosocial story themes	256	3.68	1.93	0.00	8.00	-0.43	-0.74
Narrative coherence	256	10.42	2.78	2.00	15.00	-0.37	-0.24
Behavioural dysregulation	256	6.15	3.73	0.00	16.00	0.62	-0.23

CBCL = Child Behavior Checklist; PPACS = Preschool Parental Account of Childhood Symptoms; SDQ = Strengths and Difficulties Questionnaire; PHQ-9 = Patient Health Questionnaire-9

4.4.1.3 Stability of parent-rated child behaviour and parent-rated discipline practices over time

Correlational analyses were run to explore the stability of parent-rated child behaviour and parental discipline over a two-year period. Child behaviour problems were measured using the CBCL and PPACS. Children's prosocial behaviours were measured using a subscale from the SDQ. Parental overreactive and lax discipline were measured using the Parenting Scale. As shown in Table 23, there were medium to large positive correlations between child behaviour at timepoint one and timepoint two, as measured using the CBCL (r = .54, p < .001) and PPACS (r = .39, p < .001). Children's prosocial behaviour across the two timepoints was also found to have a medium positive association, as measured through the SDQ prosocial subscale (r = .41, p < .001). There was a large, positive correlation between parental overreactivity measured across the two timepoints (r = .50, p <.001) and a small, positive correlation between parental laxness (r = .22, p < .001). This suggests there is some stability in child behaviour and parenting practices over time.

Table 23

Bivariate Correlations between Measures of Child Behaviour and Parenting Practices Measured at Baseline and 24-Months

Variable	1	2	3	4	5	6	7	8	9	10
1. PPACS total BL	-									
2. PPACS total 24M	.39**	-								
3. CBCL total BL	.46**	.41**	-							
4. CBCL total 24M	.21**	.59**	.54**	-						
5. SDQ prosocial BL	-.13*	-.09	-.06	-.06	-					
6. SDQ prosocial 24M	-.07	-.28**	-.16**	-.40**	.41**	-				
7. Parenting scale overreactivity BL	.04	.11	-.13*	.01	-.06	.02	-			
8. Parenting scale overreactivity 24M	.08	.06	-.03	-.01	-.02	.06	.50**	-		
9. Parenting scale laxness BL	-.04	.03	.01	-.05	.10	.02	-.05	.06	-	
10. Parenting scale laxness 24M	.06	.05	.03	.02	-.02	-.04	-.05	-.02	.22**	-

Note. * $p < .05$, ** $p < .01$

PPACS = Preschool Parental Account of Childhood Symptoms; CBCL = Child Behavior Checklist; SDQ = Strengths and Difficulties Questionnaire; BL = baseline assessment; 24M = 24-month follow-up assessment

4.4.1.4 Associations between baseline reports of child behaviour and parental discipline and children's outcomes on the story stem assessment two years later

Correlational analyses were run to explore associations between parent-reported child behaviour and parenting, measured at 12 to 36 months, and children's responses to the story stem assessment, measured two years later at three to five years old (see Table 24).

There were small, positive correlations between baseline scores of children's early behaviour problems and later disruptive story stem themes when early child behaviour was measured through both the CBCL (r = .22, p = <.001) and the PPACS (r = .17, p = .008). That is, higher problem scores on the parent-reported measures at baseline were associated with higher levels of disruptive themes featured in children's stories. There was also a small, positive correlation between baseline prosociality, as measured by the SDQ subscale, and prosocial story stem themes (r = .18, p = .004). That is, higher scores of early parent-reported prosocial behaviours were associated with higher levels of prosocial themes featured in children's stories. There was a small positive association between baseline prosocial behaviour and children's later narrative coherence (r = .24, p < .001), but no such associations with children's early behaviour problems (rs .06-.09). There were small, positive associations between baseline behaviour problems and children's prosocial story stem themes (CBCL: r = .16, p = .01; PPACS: r = .14, p = .028). That is, higher scores of parent-reported problem behaviours were associated with higher levels of prosocial themes two years later. There was a small, negative association between early prosociality and children's later observed dysregulation during story completion (r = -.14, p = .03), but no such associations with children's early behaviour problems (rs .04-.06).

There were small, positive associations between parental laxness and children's later prosocial themes (r = .16, p = .010) and narrative coherence (r = .15, p = .018). Thus, higher laxness (parents with a tendency to give in or not enforce rules) was associated with higher prosociality and narrative coherence two years later. Lax parenting was not significantly associated with disruptive behaviours or behavioural dysregulation. An overreactive discipline style was not significantly associated with any story stem constructs (rs -.01-.07).

Table 24

Bivariate Correlations between Baseline Parent-Report Measures of Children's Behaviour and Parenting and Child Outcomes at 24-Month Follow-Up

Variable	1	2	3	4	5	6	7	8	9
1. CBCL – total	-								
2. PPACS – total	.46**	-							
3. SDQ – prosocial	-.06	-.13*	-						
4. Parenting – laxness	.01	-.04	.10	-					
5. Parenting – over-reactivity	-.13*	.04	-.06	-.05	-				
6. Disruptive story themes	.22**	.17**	.04	-.09	-.01	-			
7. Prosocial story themes	.16*	.14*	.18**	.16**	-.04	-.15*	-		
8. Narrative coherence	.06	.09	.24**	.15*	-.06	-.19**	.68**	-	
9. Behavioural dysregulation	.06	.04	-.14*	-.04	.08	.37**	-.31**	-.36	-

Note. * $p < .05$, ** $p < .01$

CBCL = Child Behavior Checklist; PPACS = Preschool Parental Account of Childhood Symptoms; SDQ = Strengths and Difficulties Questionnaire

4.4.2 Main analysis

The overarching aim of the main analysis was to explore whether early child behaviour and early experiences of parenting were associated with children's narrative representations two years later. To explore these research questions, the analysis was split into two parts.

First, multiple linear regressions were run to explore associations between baseline parental discipline and child behaviour (entered into the same model, in separate steps), when children were 12-36 months old, and the following outcomes measured using the story stem battery two years later, when children were three to five years old:

1. Disruptive story themes
2. Behavioural dysregulation
3. Prosocial story themes
4. Narrative coherence

Separate regression models were constructed for each story stem outcome. Effect sizes were calculated for each of the regression models using Cohen's f^2, where 0.02 = small, 0.15 = medium, and 0.35 = large effects (Cohen, 1988).

Second, independent samples *t*-tests were run to explore mean group differences on each of the story stem outcomes according to group allocation (VIPP-SD or usual care)

4.4.2.1 Checking the assumptions of multiple linear regressions

Prior to conducting the regression analyses, preliminary tests were run to ensure the data met the assumptions of a hierarchical regression analysis. The findings of these checks and tests are summarised here for all regression models included in this analysis. Under the central limit theorem, we can assume with the size of this sample ($N = 256$) that our betas are from a normally distributed sampling distribution (Field, 2013). A rough rule of thumb for adequate sample size is that we would need 10-15 cases of data per predictor included in the model (Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996). For this sample, with 7-9 predictors, we would need 105-135 participants, which the current sample size surpasses.

The data was firstly examined to explore associations between included variables using correlation matrices. No variables were highly associated (above 0.8 or 0.9; Field, 2013). The variance inflation factor (VIF) and tolerance statistic, which explore whether the included predictors have a strong linear relationship with the other predictors, were also used to assess multicollinearity. All VIFs for all predictors were below 10 (Myers, 1990) and all

tolerance statistics were above 0.2 (Menard, 1995), meaning there was no evidence of multicollinearity.

Data was checked for outliers, normality, non-linearity, heteroscedasticity and independence of residuals using diagnostic statistics and by inspecting normality probability plots (P-P plots) and scatterplots.

First, plots of standardised residuals against standardised predicted values were examined. Dots appeared to be randomly and evenly dispersed throughout the plots, indicating that the assumptions of linearity and homoscedasticity were met. All standardised residuals fell between -3.30 and 3.30 apart from one case in the disruptive themes model which was 3.54 (Tabachnick & Fidell, 2014). Further tests for outliers were conducted. Cook's distance measures the overall influence of a case on the model, which values greater than 1 being a cause for concern (Cook & Weisberg, 1982). No values were greater than 1 in these models. Leverage statistics (or hat values) gauge the influence of the observed value of the dependent variable over the independent variables (Field, 2013). These values can fall between 0 and 1. Using the number of predictors and sample size, cut-off points were calculated to identify any cases having undue influence. No cases exceeded three times the average leverage value (Stevens, 2002). Finally, Mahalanobis distances were used to detect multivariate outliers. This test measures the influence of a case by examining the distance of cases from the means of the predictor variables. All values fell below the critical values for Mahalanobis distances of each model.

Partial plots were also examined to detect outliers and heteroscedasticity. There were no obvious outliers, and the clouds of dots were evenly spaced around the line, indicating homoscedasticity. To test the normality of residuals, histograms and normal probability plots were inspected. Distributions appeared to be normal for each included regression, with the histograms being fairly symmetrical and bell-shaped. The P-P plots show deviations from normality through deviations from the diagonal line. For the regressions exploring prosocial themes, narrative coherence, and observed regulation, the dots lay close to the diagonal line indicating a normal distribution. For disruptive themes, the dots mostly stayed on the diagonal line, although there was a slight deviation from the line indicating it may be non-normal.

The assumption of independence of errors was assessed using the Durbin-Watson test, which tests whether adjacent residuals are correlated. Values for this statistic vary between 0 and 4, with a value of 2 meaning the residuals are uncorrelated (Field, 2013). The Durbin-

Watson statistic for the data were: disruptive themes: 1.83, behavioural dysregulation: 1.95, prosocial themes: 1.91, narrative coherence: 2.01 suggesting that the assumption was met.

This data generally met the assumptions for a multiple linear regression.

4.4.3 Does early parental discipline and child behaviour predict children's later expression of disruptive story themes?

First, a multiple regression was conducted to explore whether early parental discipline style, measured using the Parenting Scale, and early child behaviour problems, measured using the CBCL, assessed through parent report when the children were 12 to 36 months old, predicted children's expression of disruptive themes in their stories when they were three and four years old.

Total score of disruptive themes was the dependent variable. Child age and gender, parental educational level and depressive symptoms, and intervention arm (VIPP-SD or usual care) were entered in the first step as covariates. Overreactive and lax parental discipline subscale scores were entered second. Child behaviour problems, as measured by the CBCL total score, was entered last. Table 25 displays the results of the analysis.

After controlling for potential covariates, higher levels of overreactive (standardised β = -.02, p = .71) and lax (standardised β = -.08, p = .22) parental discipline at baseline were not associated with children's expression of disruptive themes two years later, with the addition of these variables only accounting for .01 (1%) of the model's variance. Greater child behaviour problems at 12 to 36 months were associated with children's later expression of disruptive themes (standardised β = .20, p = .006), accounting for .03 (3%) of the model's variance. The overall model fit was R^2 = 0.16 (accounting for 16% of the variance). This model represents a medium effect size, f^2 = 0.19.

A second model was run using the PPACS total score as the baseline parent-reported measure of behaviour problems, rather than the CBCL. After controlling for the same variables as in the first model, and entering parental discipline in the second step, higher levels of child behaviour problems at 12-36 months were associated with higher disruptive behaviours in children's stories two years later (standardised β = .11, p = .08), accounting for .01 (1%) of the model's variance. The overall model fit was R^2 = 0.14 (accounting for 14% of the variance). This model represents a medium effect size, f^2 = 0.17. See Appendix M for the full output of this analysis.

Table 25

Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Disruptive Themes in Children's Stories Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.22 (-2.62, 7.07)	2.46		0.90	.37
	Child gender	2.53 (1.57, 3.49)	0.49	.31	5.18	.000
	Child age	0.03 (-0.04, 0.10)	0.04	.06	0.97	.34
	Caregiver education	-0.82 (-1.58, -0.07)	0.38	-.13	-2.15	.03
	Caregiver mood	0.07 (-0.06, 0.21)	0.07	.06	1.05	.30
	Trial arm	-0.07 (-1.03, 0.89)	0.49	-.01	-0.15	.88
2	(Constant)	5.59 (-2.57, 13.75)	4.14		1.35	.18
	Child gender	2.47 (1.50, 3.43)	0.49	.30	5.03	.000
	Child age	0.04 (-0.03, 0.11)	0.04	.07	1.04	.30
	Caregiver education	-0.84 (-1.60, -0.07)	0.39	-.13	-2.16	.03
	Caregiver mood	0.07 (-0.07, 0.20)	0.07	.06	1.00	.32
	Trial arm	-0.07 (-1.03, 0.89)	0.49	-.01	-0.14	.89
	Overreactive parenting	-0.20 (-1.29, 0.88)	0.55	-.02	-0.37	.71
	Lax parenting	-0.71 (-1.82, 0.41)	0.57	-.08	-1.24	.22
3	(Constant)	3.77 (-4.38, 11.92)	4.14		0.91	.36
	Child gender	2.47 (1.52, 3.43)	0.48	.30	5.11	.000
	Child age	0.02 (-0.05, 0.09)	0.04	.03	0.52	.60
	Caregiver education	-0.63 (-1.39, 0.14)	0.39	-.10	-1.61	.11
	Caregiver mood	-0.02 (-0.17, 0.13)	0.08	-.02	-0.29	.77
	Trial arm	0.04 (-0.91, 0.99)	0.48	.01	0.08	.94
	Overreactive parenting	-0.06 (-1.14, 1.01)	0.55	-.01	-0.11	.91
	Lax parenting	-0.67 (-1.77, 0.43)	0.56	-.07	-1.20	.23
	Child behaviour	0.04 (0.01, 0.07)	0.01	.20	2.79	.006

Note. R^2 = .13 for Step 1, ΔR^2 = .01 for Step 2 (*p*s = .44), ΔR^2 = .03 for Step 3 (*p*s = .006)

Caregiver mood = Patient Health Questionnaire-9 (PHQ-9) total score at baseline

Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline

Child behaviour = Child Behavior Checklist (CBCL) total score at baseline

4.4.4 Does early parental discipline and child behaviour predict children's later observed behaviour dysregulation during the completion of a story stem task?

Second, a multiple regression was conducted to explore whether early parental discipline style, measured using the Parenting Scale, and early child behaviour problems, measured using the CBCL, assessed through parent report when the children were 12 to 36 months old, predicted children's ability to regulate their behaviour during the completion of the story stem task when they were three and four years old.

Total score of observed behavioural dysregulation was the dependent variable. Child age and gender, parental educational level and depressive symptoms, and intervention arm (VIPP-SD or usual care) were entered in the first step as covariates. Overreactive and lax parental discipline subscale scores were entered second. Child behaviour problems, as measured by the CBCL total score, was entered last. Table 26 displays the results of the analysis.

After controlling for potential covariates, higher levels of overreactive (standardised β = .04, p = .51) and lax (standardised β = .01, p = .91) parental discipline at baseline were not associated with children's observed dysregulation two years later, with the addition of these variables only accounting for .002 (0.2%) of the model's variance. Greater child behaviour problems at 12 to 36 months were also not associated with children's observed behavioural dysregulation (standardised β = .07, p = .31), accounting for .004 (0.4%) of the model's variance. The overall model fit was R^2 = 0.16 (accounting for 16% of the variance). This model represents a medium effect size, f^2 = 0.19.

A second model was run using the PPACS total score as the baseline parent-reported measure of behaviour problems, rather than the CBCL. After controlling for the same variables as in the first model, and entering parental discipline in the second step, higher levels of child behaviour problems at 12 to 36 months were not associated with higher observed dysregulation two years later (standardised β = -.01, p = .84), accounting for .000 of the model's variance. The overall model fit was R^2 = 0.16 (accounting for 16% of the variance). This model represents a medium effect size, f^2 = 0.17. See Appendix N for the full output of this analysis.

Table 26

Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Observed Behavioural Dysregulation Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	6.40 (2.08, 10.72)	2.19		2.92	.004
	Child gender	2.75 (1.89, 3.61)	0.44	.37	6.33	.000
	Child age	-0.05 (-0.11, 0.02)	0.03	-.09	-1.47	.14
	Caregiver education	-0.58 (-1.26, 0.09)	0.34	-.10	-1.71	.09
	Caregiver mood	0.01 (-0.11, 0.13)	0.06	.01	0.23	.82
	Trial arm	-0.39 (-1.25, 0.46)	0.44	-.05	-0.91	.37
2	(Constant)	4.61 (-2.68, 11.90)	3.70		1.25	.21
	Child gender	2.75 (1.89, 3.61)	0.44	.37	6.27	.000
	Child age	-0.04 (-0.11, 0.02)	0.03	-.08	-1.32	.19
	Caregiver education	-0.55 (-1.23, 0.13)	0.35	-.10	-1.59	.11
	Caregiver mood	0.02 (-0.10, 0.14)	0.06	.02	0.28	.78
	Trial arm	-0.38 (-1.24, 0.48)	0.44	-.05	-0.87	.39
	Overreactive parenting	0.32 (-0.65, 1.29)	0.50	.04	0.66	.51
	Lax parenting	0.06 (-0.94, 1.06)	0.51	.01	0.12	.91
3	(Constant)	4.01 (-3.37, 11.39)	3.75		1.07	.29
	Child gender	2.75 (1.89, 3.62)	0.44	.37	6.28	.000
	Child age	-0.05 (-0.11, 0.02)	0.03	-.09	-1.49	.14
	Caregiver education	-0.48 (-1.17, 0.21)	0.35	-.08	-1.36	.17
	Caregiver mood	-0.01 (-0.15, 0.12)	0.07	-.01	-0.19	.85
	Trial arm	-0.34 (-1.21, 0.52)	0.44	-.05	-0.79	.43
	Overreactive parenting	0.37 (-0.60, 1.34)	0.49	.05	0.75	.45
	Lax parenting	0.07 (-0.93, 1.07)	0.51	.01	0.14	.89
	Child behaviour	0.01 (-0.01, 0.04)	0.01	.07	1.03	.31

Note. R^2 = .16 for Step 1, ΔR^2 = .002 for Step 2 (*p*s = .80), ΔR^2 = .004 for Step 3 (*p*s = .31)

Caregiver mood = Patient Health Questionnaire-9 (PHQ-9) total score at baseline

Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline

Child behaviour = Child Behavior Checklist (CBCL) total score at baseline

4.4.5 Does early parental discipline and child behaviour (prosociality and behaviour problems) predict children's later expression of prosocial story themes?

Third, a multiple regression was conducted to explore whether parent-reported early parental discipline style (measured using the Parenting Scale), early child behaviour problems (measured using the CBCL), and early child prosociality (measured using the SDQ prosocial subscale), assessed when children were 12 to 36 months old, predicted children's expression of prosocial story themes at three and four years old.

Total score of prosociality in children's story stem responses was the dependent variable. Child age and gender, parental educational level and depressive symptoms, and intervention arm (VIPP-SD or usual care) were entered in the first step as covariates. Overreactive and lax parental discipline scale scores were entered second. Child behaviour problems, as measured by the CBCL total score, was entered third. Child prosociality, as measured using the SDQ subscale score, was entered last. Table 27 displays the results of the analysis.

After controlling for potential covariates, higher levels of overreactive (standardised β = .05, p = .38) and lax (standardised β = .08, p = .14) parental discipline at baseline were not associated with children's expression of prosocial themes two years later, with the addition of these variables accounting for .01 (1%) of the model's variance. Greater child behaviour problems at 12 to 36 months were also not associated with children's later expression of prosocial themes (standardised β = .07, p = .30), accounting for .003 (0.3%) of the model's variance. Greater prosociality at 12-36 months was also not associated with children's later expression of prosocial themes (standardised β = .01, p = .84), accounting for .000 of the model's variance. The overall model fit was R^2 = 0.29 (accounting for 29% of the variance). This model represents a large effect size, f^2 = 0.40.

Table 27

Linear Model of Early Parental Discipline and Child Behaviour as Predictors of Children's Expression of Prosocial Story Themes Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	0.69 (-1.39, 2.78)	1.06		0.65	.51
	Child gender	-1.21 (-1.62, -0.80)	0.21	-.31	-5.78	.000
	Child age	0.09 (0.06, 0.12)	0.02	.33	6.01	.000
	Caregiver education	-0.02 (-0.34, 0.30)	0.16	-.01	-0.13	.90
	Caregiver mood	0.04 (-0.02, 0.09)	0.03	.07	1.22	.23
	Trial arm	0.89 (0.48, 1.30)	0.21	.23	4.26	.000
2	(Constant)	-1.53 (-5.02, 1.97)	1.78		-0.86	.39
	Child gender	-1.18 (-1.59, -0.77)	0.21	-.31	-5.62	.000
	Child age	0.09 (0.06, 0.12)	0.02	.33	5.82	.000
	Caregiver education	-0.00 (-0.33, 0.32)	0.17	-.00	-0.02	.99
	Caregiver mood	0.04 (-0.02, 0.10)	0.03	.07	1.31	.19
	Trial arm	0.89 (0.48, 1.31)	0.21	.23	4.27	.000
	Overreactive parenting	0.21 (-0.26, 0.67)	0.24	.05	0.88	.38
	Lax parenting	0.36 (-0.12, 0.83)	0.24	.08	1.47	.14
3	(Constant)	-1.80 (-5.34, 1.74)	1.80		-1.00	.32
	Child gender	-1.18 (-1.59, -0.77)	0.21	-.31	-5.62	.000
	Child age	0.09 (0.06, 0.12)	0.02	.32	5.52	.000
	Caregiver education	0.03 (-0.30, 0.36)	0.17	.01	0.18	.86
	Caregiver mood	0.02 (-0.04, 0.09)	0.03	.05	0.73	.47
	Trial arm	0.91 (0.50, 1.32)	0.21	.24	4.34	.000
	Overreactive parenting	0.23 (-0.24, 0.69)	0.24	.05	0.97	.33
	Lax parenting	0.36 (-0.12, 0.84)	0.24	.08	1.49	.14
	Child behaviour	0.01 (-0.01, 0.02)	0.01	.07	1.02	.31
4	(Constant)	-1.82 (-5.37, 1.73)	1.80		-1.01	.31
	Child gender	-1.17 (-1.59, -0.75)	0.21	-.30	-5.46	.000
	Child age	0.09 (0.05, 0.12)	0.02	.31	5.08	.000

Variable	*B* (95% CI)	SE B	β	*t*	*p*
Caregiver education	0.03 (-0.30, 0.36)	0.169	.01	0.18	.86
Caregiver mood	0.02 (-0.04, 0.09)	0.033	.05	0.73	.47
Trial arm	0.91 (0.49, 1.32)	0.210	.24	4.32	.000
Overreactive parenting	0.23 (-0.24, 0.70)	0.236	.05	0.97	.33
Lax parenting	0.36 (-0.12, 0.84)	0.243	.08	1.48	.14
Child behaviour	0.01 (-0.01, 0.02)	0.006	.07	1.04	.30
Child prosociality	0.01 (-0.09, 0.11)	0.051	.01	0.20	.84

Note. R^2 = .28 for Step 1, ΔR^2 = .01 for Step 2 (*p*s = .24), ΔR^2 = .003 for Step 3 (*p*s = .31), ΔR^2 = .000 for Step 4 (*p*s = .84)

Caregiver mood = Patient Health Questionnaire-9 (PHQ-9) total score at baseline

Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline

Child behaviour = Child Behavior Checklist (CBCL) total score at baseline

Child prosociality = Strengths and Difficulties Questionnaire (SDQ) subscale score at baseline

4.4.6 Does early parental discipline and child behaviour (prosociality and behaviour problems) predict children's later narrative coherence?

Fourth, a multiple regression was conducted to explore whether parent-reported early parental discipline style (measured using the Parenting Scale), early child behaviour problems (measured using the CBCL), and early child prosociality (measured using the SDQ prosocial subscale), assessed when children were 12 to 36 months old, predicted children's later narrative coherence during their storytelling at three and four years old.

Total score of narrative coherence was the dependent variable. Child age and gender, parental educational level and depressive symptoms, and intervention arm (VIPP-SD or usual care) were entered in the first step as covariates. Overreactive and lax parental discipline scale scores were entered second. Child behaviour problems, as measured by the CBCL total score, was entered third. Child prosociality, as measured using the SDQ subscale score, was entered last. Table 28 displays the results of the analysis.

After controlling for potential covariates, higher levels of overreactive (standardised β = 0.03, p = .59) and lax (standardised β = .07, p = .23) parental discipline at baseline were not associated with children's narrative coherence two years later, with the addition of these variables only accounting for .01 (1%) of the model's variance. Greater child behaviour problems at 12 to 36 months were also not associated with children's later narrative coherence (standardised β = -.05, p = .41), accounting for .002 (0.2%) of the model's variance. Greater prosociality at 12-36 months was also not associated with children's narrative coherence (standardised β = .07, p = .25), accounting for .004 (0.4%) of the model's variance. The overall model fit was R^2 = 0.27 (accounting for 27% of the variance). This model represents a large effect size, f^2 = 0.37.

Table 28

Linear Model of Early Parental Discipline and Child Behaviour as Predictors of Children's Narrative Coherence Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	4.93 (1.91, 7.95)	1.53		3.21	.001
	Child gender	-1.78 (-2.38, -1.19)	0.30	-.32	-5.90	.000
	Child age	0.13 (0.09, 0.17)	0.02	.33	5.93	.000
	Caregiver education	0.37 (-0.10, 0.84)	0.24	.09	1.57	.12
	Caregiver mood	0.05 (-0.04, 0.13)	0.04	.06	1.14	.26
	Trial arm	1.07 (0.47, 1.66)	0.30	.19	3.52	.001
2	(Constant)	2.62 (-2.45, 7.69)	2.58		1.02	.31
	Child gender	-1.75 (-2.35, -1.15)	0.30	-.32	-5.75	.000
	Child age	0.13 (0.08, 0.17)	0.02	.33	5.69	.000
	Caregiver education	0.39 (-0.09, 0.86)	0.24	.09	1.61	.11
	Caregiver mood	0.05 (-0.03, 0.14)	0.04	.07	1.19	.23
	Trial arm	1.07 (0.47, 1.66)	0.30	.19	3.52	.001
	Overreactive parenting	0.18 (-0.49, 0.85)	0.34	.03	0.53	.59
	Lax parenting	0.42 (-0.27, 1.11)	0.35	.07	1.20	.23
3	(Constant)	2.94 (-2.19, 8.07)	2.61		1.13	.26
	Child gender	-1.75 (-2.35, 1.15)	0.31	-.32	-5.74	.000
	Child age	0.13 (0.09, 0.18)	0.02	.34	5.74	.000
	Caregiver education	0.35 (-0.13, 0.83)	0.25	.08	1.43	.15
	Caregiver mood	0.07 (-0.03, 0.16)	0.05	.09	1.43	.15
	Trial arm	1.05 (0.45, 1.65)	0.30	.19	3.44	.001
	Overreactive parenting	0.16 (-0.52, 0.83)	0.34	.03	0.46	.65
	Lax parenting	0.41 (-0.28, 1.11)	0.35	.07	1.18	.24
	Child behaviour	-0.01 (-0.02, 0.01)	0.01	-.05	-0.82	.41
4	(Constant)	2.78 (-2.36, 7.92)	2.61		1.07	.29
	Child gender	-1.68 (-2.29, -1.07)	0.31	-.30	-5.40	.000
	Child age	0.12 (0.07, 0.17)	0.02	.31	4.97	.000

Variable	*B* (95% CI)	SE B	β	*t*	*p*
Caregiver education	0.36 (-0.13, 0.84)	0.25	.08	1.46	.15
Caregiver mood	0.07 (-0.03, 0.16)	0.05	.09	1.41	.16
Trial arm	1.04 (0.44, 1.64)	0.30	.19	3.41	.001
Overreactive parenting	0.16 (-0.51, 0.84)	0.34	.03	0.47	.64
Lax parenting	0.40 (-0.29, 1.09)	0.35	.06	1.13	.26
Child behaviour	-0.01 (-0.02, 0.01)	0.01	-.04	-0.62	.54
Child prosociality	0.09 (-0.06, 0.23)	0.07	.07	1.16	.25

Note. R^2 = .26 for Step 1, ΔR^2 = .01 for Step 2 (*p*s = .43), ΔR^2 = .002 for Step 3 (*p*s = .41), ΔR^2 = .004 for Step 3 (*p*s = .25)

Caregiver mood = Patient Health Questionnaire-9 (PHQ-9) total score at baseline

Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline

Child behaviour = Child Behavior Checklist (CBCL) total score at baseline

Child prosociality = Strengths and Difficulties Questionnaire (SDQ) subscale score at baseline

4.4.7 Considering the effects of a video-feedback parenting intervention on children's story stem outcomes.

Randomised controlled trials allow us to explore differences in outcomes between groups and attribute differences to the receipt of the intervention. Independent *t*-tests were used to compare children's outcomes, as measured through the story stem assessment, by trial arm allocation (VIPP-SD vs. usual care) at two years post-randomisation.

The sample sizes of families randomised to VIPP-SD (n = 128) and UC (n = 128) were equal. The distributional assumptions necessary for an independent *t*-test are discussed in Section 4.4.2.1. Levene's test, a test of the homogeneity of variance, were non-significant for each test indicating equal variance. Standardised effect sizes (Cohen's *d*) were computed using the difference in means between the two groups and dividing this by the standard deviation of the control group (usual care). Cohen's *d* effect sizes can be interpreted using the guide of 0.20 = small, 0.50 = medium, and 0.80 = large (Cohen, 1988).

Disruptive themes. The mean score for disruptive themes in children's stories for families randomised to receive the VIPP-SD programme (M = 4.89, SE = 0.35) and for families randomised to usual care (M = 5.12, SE = 0.37) was similar. The difference between the two groups, 0.27, 95% CI [-0.79, 1.24], $t(254) = 0.44$, $p = .66$, represented a very small effect, d = -0.05.

Behavioural dysregulation. The mean score for dysregulated behaviour during children's storytelling for families randomised to receive the VIPP-SD programme (M = 5.88, SE = 0.32) and for families randomised to receive usual care (M = 6.42, SE = 0.32) was similar. The difference between the two groups, 0.55, 95% CI [-0.37, 1.46], $t(254) = 1.18$, $p = .24$, represented a very small effect, d = -0.14.

Prosocial themes. Children whose families were randomised to receive the VIPP-SD programme scored more highly on prosocial themes (M = 4.13, SE = 0.17) than children whose families were randomised to usual care (M = 3.23, SE = 0.16). This difference, -0.90, 95% CI [-1.36, -0.44], $t(254) = -3.83$, $p < .001$, represented a medium effect size, d = 0.49.

Narrative coherence. Children whose families were randomised to receive the VIPP-SD programme scored more highly on narrative coherence (M = 10.95, SE = 0.24) than children whose families were randomised to usual care (M = 9.89, SE = 0.24). This difference, -1.06, 95% CI [-1.74, -0.39], $t(254) = -3.11$, $p = .002$, represented a small effect size, d = 0.39.

4.5 Discussion

This study explored early predictors of three- and four-year-old's narrative representations of prosocial and disruptive behaviour, narrative coherence, and behavioural dysregulation during story completion. The research findings are mixed. Children's early behaviour problems, but not early parental discipline style, predicted their expression of disruptive story themes two years later. However, children's expression of prosocial themes, their narrative coherence, and their behavioural dysregulation were not predicted by early child behaviour nor early parental discipline practices. Extending further upon the investigation of the influential role of parents in the development of children's internal representations, this study also examined potential effects of a video-feedback intervention, aiming to enhance parental sensitivity and sensitive discipline, on children's narrative representations. The intervention had no effects on children's behavioural dysregulation during storytelling, or their expression of disruptive narrative themes two years later but demonstrated small effects on their narrative coherence and the expression of prosocial themes. Research findings are now discussed in more detail, before the strengths, limitations, and implications of the study are considered.

4.5.1 Summary of findings

4.5.1.1 Aim 1: Longitudinal associations between early child behaviour and story stem outcomes.

Early behaviour problems rated at 12 to 36 months predicted children's expression of disruptive themes but not prosocial themes, narrative coherence, or their capacity to regulate their behaviour during story completion at three to five years of age. Longitudinal research exploring early predictors of these domains in preschool populations is scarce and findings are mixed. Interpretation and potential explanations for each construct are considered in turn.

4.5.1.1.1 Disruptive themes

Children's behaviour problems at one to three years of age were associated with their expression of disruptive narrative themes at three to five years. This could suggest that disruptive responses to narrative prompts are reflective of enduring behavioural difficulties initially detected in toddlerhood. Challenging behaviours are a typical characteristic of toddler-aged children and usually decline from age three (Alink et al., 2006). However, even very early signs of intense or frequent behaviour problems can signal enduring difficulties for some children. Indeed, in this sample there was a moderate degree of stability in parent-

ratings of child behaviour over two years (r = .54). Taken with the findings of the previous chapter, where modest concurrent associations between parent-rated child behaviour and children's disruptive themes were found (rs .18-.24), we can assume that, rather than playful fantasy, disruptive story themes may be somewhat reflective of children's dysregulated and challenging responses to real-life interactions and experiences. Building on the mostly concurrent body of previous research, the present finding provides us with important insight into the ways in which early behavioural functioning, measured at only one and two years old, may be associated with children's own perspectives of their later functioning.

4.5.1.1.2 Prosocial themes

Children's early prosocial tendencies did not predict their later expression of prosocial narrative themes. Prosocial development is considered to be shaped by a complex interplay of psychological and dispositional processes, socialisation, and biological and contextual factors (Eisenberg, Fabes, & Spinrad, 2006). One possibility is that the two years between study assessments (occurring at 1-3 and 3-5 years) spanned a sensitive period in children's prosocial development. While some forms of prosocial behaviour develop in toddlerhood (e.g., cooperating), others often do not develop until later in early childhood (e.g., sharing, helping) (Köster, Ohmer, Nguyen, & Kärtner, 2016; Svetlova, Nichols, & Brownell, 2010). This developmental shift is considered to be driven by increases in social-cognitive understanding, emotional maturity, and self-awareness, acquired largely through socialisation (Hoffman, 2001; Paulus et al., 2015). Transitions into full-time childcare and formal education through the study's duration will have exposed this sample to new social settings and behavioural expectations that may have been formative for their prosocial development. However, such a developmental leap is unlikely to fully explain this finding given that a moderate degree of stability in parent-rated prosociality was found over two years (r = .41). Differences in the nature of the assessment of prosociality at each time point may partly explain the lack of association, including differences in informants (parents and child), measures (SDQ and story stems), and definitions of prosocial behaviour. While concurrent findings from Chapter 3 indicated that parent- and child-reported prosociality were modestly related (r = .21) it seems likely that each measure taps into differing elements of the broader construct. For instance, while caregivers reported on behaviours they observed in their children as one- and two-year-olds, children's expression of (non-) prosocial story content may reflect their experiences in other settings (e.g., school) that their caregivers have less opportunity to observe at three and four years old.

4.5.1.1.3 Narrative coherence

Early behaviour did not predict the coherence with which children told their stories. Most previous studies exploring these associations have been concurrent in design and/or included older children and adolescents. The longitudinal nature of these findings provides us with unique insight into the early developmental influences on preschool-aged children's narrative coherence. One possible explanation for the lack of association may be that preschool-aged children as a whole exhibit a general low level of coherence that diminishes our ability to identify early predictors of the construct. Associations between age and narrative coherence were reported in Chapter 3. Exploring scores by age, 26% of three-year-olds and 47% of four-year-olds told consistently coherent stories (with 'coherence' defined as an average score of 4 or above across three stories). This could indicate that, firstly, the necessary skills for a coherent story (e.g., emotional regulation, cognitive processing, planning, sequencing) are still undergoing development in this age-range as a whole and, second, that there may be a developmental shift in these abilities between three and four years. These developmental differences may cloud our ability to identify early predictors of this construct when we analyse children spanning two years as one age group. Alternatively or additionally, unmeasured relational factors (e.g., supportive parenting practices, attachment quality) could potentially 'buffer' against the ways in which behavioural difficulties may disrupt the organisation and coherence of children's internal systems, reflected in their narrative coherence (e.g., Hill et al., 2007).

4.5.1.1.4 Behavioural dysregulation

Finally, early child behaviour did not predict observations of children's behavioural dysregulation two years later. In the previous chapter, concurrent measures of externalising behaviour problems and observed dysregulation were found to be modestly associated (parent report: *r*s .17-.18; teacher report: $r = .26$). According to Kopp's (1982) model of self-regulation, children's regulatory skills undergo considerable development between the ages of one and three years. One possible explanation for the lack of longitudinal association in this chapter could be that the behavioural expectations captured through this measure at three and four years old are generally not expected of one- and two-year-olds and thus behavioural difficulties exhibited at this earlier age may have been less connected to the child's capacity for regulation. Behaviours assessed for dysregulation in this study included the child's capacity to listen to instructions, remain seated, wait their turn, and refrain from indiscriminately touching test equipment. Parallels with behavioural expectations in the

classroom could be drawn (note, the stronger concurrent association with teacher-rated behaviour). Self-regulation is considered to be shaped by the internalisation of expectations and values that children become aware of through the demands of their social environments, under the guidance of adult regulators such as parents and teachers (Karreman et al., 2006). Early behavioural difficulties may be less indicative of future challenges with regulation because of the important relational and environmental influences (e.g., exposure to childcare settings, peers, and formal education) that drive the development of regulatory skills.

4.5.1.2 Aim 2: Longitudinal associations between early parental discipline style and story stem outcomes.

Children's outcomes measured through the story stem battery were not predicted by early parental discipline style. Several factors may contribute to these findings. First, discipline behaviours measured in this study were mild, with most items focused on parents' use of verbal discipline or their ability to 'follow through' on boundaries set. More harsh or frightening discipline practices may be more disruptive to the formation and organisation of children's internal working models and thus predictive of children's narrative representations. Robust links between experiences of neglect, abuse, and maltreatment and negative aspects of internal representations have been established (Grych, Wachsmuth-Schlaefer, & Klockow, 2002; Hillman et al., 2020; Pickreign Stronach et al., 2011). Nonetheless, in the context of previous research, the importance of milder dysfunctional discipline approaches for children's functioning should not be ruled out entirely. Laible and Thompson (2002) reported that parental response to parent-child conflict (specifically, ability to resolve conflict, and use of compromise and bargaining) at 30 months predicted children's prosocial story themes and narrative coherence, but not aggressive themes, at 36 months. Mackenbach et al. (2014) reported that maternal and paternal harsh discipline at age three predicted child-reported behaviour problems measured using a puppet-based interview at age six. One factor that may contribute to the lack of such findings in the current sample could be that discipline was self-reported. Given its sensitive nature, reports from parents are likely to be subject to social desirability bias. Another possibility is that other aspects of the parent-child relationship may play a confounding role in the relationship between discipline and internal representations. For instance, McKee et al. (2007) reported that parental warmth acted as a buffer for children against the detrimental effects of harsh physical discipline.

4.5.1.3 Aim 3: Exploring the effects of a parenting intervention on children's outcomes.

To further our understanding of the early influences of parenting on children's internal representations, the experimental design of the HSHS trial was utilised. This analysis explored whether receipt of VIPP-SD, a video-feedback parenting intervention, was associated with aspects of children's narrative representations. Group differences were not found for children's expression of disruptive themes or their observed behavioural dysregulation but there were small to medium effects on children's narrative coherence and their expression of prosocial themes. Possible explanations for the effects on children's coherence and prosociality are now considered in the context of key targets of the VIPP-SD intervention.

One possibility is that improvements in child behaviour mediate the positive associations between the intervention and children's narrative outcomes. There is a robust evidence base for the effects of the VIPP-SD programme on child behaviour (Juffer, Bakermans-Kranenburg, et al., 2017b), and indeed reductions in behaviour problems were seen in this sample immediately post-intervention and at the two-year follow-up (reported in O'Farrelly, Barker, et al., 2021; O'Farrelly, Watt, et al., 2021). It is interesting to note that the group differences seen in the current study were linked to story stem outcomes with a more relational element to them. Because of this, an alternative or additional explanation is that the programme enhanced features of parenting that had positive effects on children's internal systems. A hallmark of attachment theory is that the quality and nature of children's ongoing experiences with their caregivers feeds into their internal representations, acting as a lens through which children interpret their relational worlds (Schechter et al., 2007). It is these internal representations that children are then considered to draw upon when organising and structuring their narratives (Moss et al., 2009; Sher-Censor & Oppenheim, 2004; Wong et al., 2011). Both parental sensitivity and sensitive discipline style are key targets of the VIPP-SD intervention and either could play a mediating role in the relationships seen. Further to this, encouraging caregivers to share emotions and demonstrate empathy towards their children are central themes in the VIPP-SD therapists' feedback (Juffer, Bakermans-Kranenburg, et al., 2017b). It may be that parents' engagement in more emotionally sensitive dialogue enriched the emotional coherence of the parent-child dyads. Such improvements in the daily interactions of children's lives may have offered a model for their relational skills with others, subsequently improving children's capacity to organise and structure their narratives in a prosocial way. The experimental nature of these findings provides empirical support for the

importance of parenting in the development of children's narrative representations, contributing to a scant collection of previous findings exploring the effects of parenting interventions on measures aiming to access children's worlds (e.g., Olds et al., 2004; Reese et al., 2020; Toth et al., 2002). Future studies would benefit from exploring mediating mechanisms of such relationships by investigating possible changes in parent-child interaction post-intervention.

In summary of these findings, the longitudinal and experimental nature of this study provides important insight into the formation of children's internal representations, accessed through a story stem battery. First, even very early behaviour problems, measured as young as one year old, may act as a signal for later difficulties with specific regard to children's representations of disruptive behaviour. Second, while overreactive and lax parental discipline were not found to predict children's later representations, a parenting intervention targeting positive parenting practices appeared to exert an effect on the coherence and prosociality of children's later narrative representations. This suggests that possible improvements in parenting may have important effects on the organisation and expression of children's internal representations.

4.5.2 Strengths and limitations

Here I reflect on the strengths and limitations of the study's sample and design, as well as its measures of children's internal representations, child behaviour, and parental discipline.

4.5.2.1 Sample and design

There are relatively few studies using child informant measures with preschool-aged children, and fewer still of preschool-aged children recruited based on early reports of behavioural difficulties. Given the long-term outcomes associated with enduring behaviour problems, this sample offers an important population to study during a significant stage of behavioural development where problems are not yet embedded. The longitudinal, experimental design of the study allows for an inquiry into possible developmental risk or protective factors in the shaping of components of children's internal representations. Such understanding is important given that children's internal representations are considered to contribute to multiple elements of their functioning (Woolgar, 1999). This study's design also benefits from the high retention rate achieved over two years. Sufficient data for inclusion in the current study's analysis was available for 85% of the baseline sample which far surpasses

the average retention of 68% reported for follow-up studies involving young children (Karlson & Rapoff, 2009).

However, the extent to which these findings can be generalised to populations beyond that studied is less known. We know there are differences in the narratives of children with and without behavioural problems (Beresford, Robinson, Holmberg, & Ross, 2007; Zahn-Waxler et al., 1994). It could be that factors that may be more protective or predictive of disorganised or negative representations in a general population were not so clear to see in this sample due to a lower level of variability in children's behaviour. We also cannot be sure that patterns of association would be the same in a clinically-diagnosed population. Accurately identifying behaviour problems in one- and two-year-olds is challenging due to, first, the typical developmental trajectories of these difficulties (peaking at age two and then declining for most children), and second, a lack of robust, well-validated screening tools for this age-range. While there was a moderate level of stability in behaviour problems across the two years of the study, indicating a level of success in identifying enduring difficulties, we are still lacking in an understanding of what the current findings may mean for children with a clinical diagnosis of a behavioural disorder. Further to this, considering the findings of previous research, associations may be different in populations recruited due to parental- or environmental-risk, such as children at risk of maltreatment or neglect (e.g., Olds et al., 2004; Toth et al., 1997) or adopted children (Hillman et al., 2020).

4.5.2.2 Measurement of variables

4.5.2.2.1 Story stems

A key strength of this study is its inclusion of a child-informant measure with children aged only three and four. Such measures are under-utilised in both research and clinical practice, especially with children so young. The findings of the previous chapter provided a level of confidence in the measure's use and utility with this sample. Extending upon this, its inclusion in the current study suggests that unique insight can be gained from such measures of children's narratives and that they may provide a helpful complement to adult-reported measures of child development. However, given the novelty of key aspects of the measure's administration and analysis, further developments and investigations into its use with other populations may be valuable.

4.5.2.2.2 Child behaviour

Child behaviour was measured using a researcher-led clinical interview (PPACS) and two adult-rated questionnaires (CBCL and SDQ prosocial subscale). As detailed in the method, these measures have exhibited promising psychometric properties when utilised with children aged under five (e.g., CBCL: Ivanova et al., 2010; Rescorla et al., 2011; PPACS: Sonuga-Barke et al., 1994; Taylor et al., 1986a; SDQ prosocial subscale: Dahlberg et al., 2019; Gustafsson et al., 2016). However, study of these properties with children aged under three years still remains a relatively understudied area, and further investigations would increase our confidence in the validity of findings. Further to this, there are other child-related variables, other than behaviour, that may have had notable effects on children's narrative content and structure but that were unmeasured in this study. For instance, child temperament (Goldwyn, Stanley, Smith, & Green, 2000), executive functions (Dealy et al., 2019), and language comprehension and production (Dealy et al., 2019; Fiorentino & Howe, 2004; Torng & Sah, 2020) have been linked to children's use of prosocial and aggressive themes and their narrative coherence. For the current sample, language ability seems like an especially important missing component given that behaviour problems are often accompanied by language delay or impairment (van Daal, Verhoeven, & van Balkom, 2007).

4.5.2.2.3 Parenting behaviours

The concept of 'parenting' is broad and formed out of a complex web of parental behaviours and traits. According to attachment theory, the quality of relationships and interactions with key attachment figures are internalised by children and integrated into their internal working models (Bretherton, Ridgeway, et al., 1990; Laible et al., 2004). In the current study, early parenting was only directly assessed through a measure of lax and overreactive discipline practices. Other components of parenting and the parent-child relationship may be just as, if not more, important in the formation and expression of children's internal representations. Previous studies of children's narratives have recognised the contributions of parental warmth (Laible et al., 2004), parental sensitivity (Bretherton, Biringen, Ridgeway, Maslin, & Sherman, 1989), parental communication style (Reese et al., 2020), parental violence (Minze, McDonald, Rosentraub, & Jouriles, 2010), the triadic relationship (von Klitzing & Bürgin, 2005), and, especially, attachment security (Bretherton, Prentiss, et al., 1990; Oppenheim, 1997; Sher-Censor & Oppenheim, 2004; Wong et al., 2011) on children's internal representations. Further study of these additional components of parenting may enhance our insight into the mechanisms underlying the current findings.

4.5.3 Implications and conclusions

Building on the investigations of its use in the previous chapter, this study indicates that a play-based story stem narrative assessment may offer a useful tool to explore the longitudinal influences of children's early experiences on elements of their internal representations. Such measures have the potential to help us learn about the internal thought processes, feelings, and experiences of very young children during a key developmental stage. These findings are in the minority of research exploring such associations in children so young who were recruited based on early reports of behaviour problems. We gain further insight into the potential long-terms effects of early behaviour problems on children's internal representations, specifically regarding their expressions of disruptive behaviours. This reiterates the importance of developing effective supports for families at the earliest signs of these difficulties. These findings also offer partial support for the view that parenting is an important contributing factor to the formation of children's internal representations. While parental discipline was not associated with children's narrative representations, the exploration of the effects of a video-feedback parenting intervention indicate that children's narrative coherence and prosocial themes may be associated with positive parenting practices (i.e., the targets of the VIPP-SD programme). Further interrogation of the precise mechanisms underlying these effects is necessary. However, based on previous theoretical and empirical findings in this field, one possibility is that this parenting intervention improved aspects of the parent-child relationship (e.g., the quality of the attachment relationship) that are reflected in children's narratives. It seems reasonable to suggest that parenting interventions may not only have the potential to improve parenting and children's external behaviours, as has been reported in other studies, but also the internal worlds of children.

CHAPTER 5: The place of sensitivity: Examining associations between parental sensitivity, behaviour problems, and children's narrative representations in the context of a video-feedback parenting intervention

5.1 Chapter rationale

Attachment theory proposes that the construction of children's internal representations is influenced by early interactions with those around them including, of particular importance, their immediate caregivers. This chapter builds on the findings of the previous chapter by exploring the potential mechanisms behind some of the associations found. The attention remains on the early influences of parenting, with a specific focus on parental sensitivity – a central construct in the study of parent-child interactions. First, associations between parental sensitivity in toddlerhood and domains of children's narrative representations at three and four years old are examined. Following this, I explore whether the effects of the VIPP-SD parenting programme on children's narrative coherence, reported in Chapter 4, are mediated by improvements to parental sensitivity. Finally, I test the effects of VIPP-SD on children's behaviour problems before examining whether parental sensitivity post-intervention mediates any changes in children's behaviour.

Note. Due to the resource-intensive nature of conducting observations of parental sensitivity, the population involved in the forthcoming analysis is a subsample (~50%) of the full Healthy Start, Happy Start sample. The limitations and implications of this reduced sample size are discussed throughout the chapter.

5.2 Introduction

A central proposal of attachment theory is that the quality of children's early caregiving experiences become internalised and generalised into a set of expectations about their attachment relationships, shaping social behaviour more generally (Society for Emotion and Attachment Studies, 2021). These internal expectations are not considered to be fixed but rather receptive to new interactions and experiences. In the previous chapter, children's experiences of lax and overreactive parental discipline were not found to be associated with features of their story stem narratives, considered to evoke constructs related to children's inner systems. The current chapter extends this focus on parenting by exploring associations between parental sensitivity and children's internal and external functioning.

In this section, I focus closely on concepts related to sensitive and responsive caregiving. Specifically, I consider ideas around the definition, measurement, influence, and promotion of parental sensitivity before reflecting upon its proposed links to children's external functioning and internal representations, measured through their narratives.

5.2.1 Parenting and children's development

A consistent body of research indicates that the paths of vulnerability and resilience to common mental health problems begin to be laid in the first few years of children's lives (Pine & Fox, 2015; Wakschlag et al., 2019). Thus, early childhood provides us with a critical opportunity to support and promote positive child outcomes, not only to benefit children and families, but also the multiple services that often feel the ripple effects of children's enduring challenges, such as the health, education, and social care systems (Britto et al., 2017; Costello & Maughan, 2015).

Robust evidence indicates that early family relationships and interactions are an especially salient component in the onset and maintenance of children's earliest behavioural difficulties, making these a critical target of intervention (Miner & Clarke-Stewart, 2008; Pinquart, 2017a, 2017b). For infants and toddlers, their immediate caregivers are a primary source of these early interactions and familial bonds. Thus, parents (i.e., any caregivers with primary responsibility of a child) can fulfil an exceptionally unique, powerful, and consequential role in their children's lives as one of the most central influences in their development. Key to supporting parents as they navigate this role is a clear understanding of the kinds of parenting behaviours and styles that are most beneficial for children's development.

Defining and conceptualising 'positive' and 'negative' parenting has formed the focus of decades of research covering a vast expanse of concepts including the quality (e.g., reciprocity, playfulness, warmth, sensitivity), content (e.g., semantic meaning of verbalisations), and frequency (e.g., questions, vocalisations) of varied parent-related behaviours (Aspland & Gardner, 2003). Research into 'positive' parenting has largely focused on parental sensitivity, responsiveness, and warmth, while investigations of 'negative' parenting have frequently studied negative control and intrusive, neglectful, hostile, and permissive parenting (Oliver & Pike, 2018; Pinquart, 2017a, 2017b). Here, I focus in on the concept of sensitive caregiving.

5.2.2 Sensitive caregiving

Sensitive and responsive caregiving are regarded as centrally important features of a positive parent-child relationship to promote children's happy and healthy development. In the first few years of life, children's needs are entirely dependent on the capabilities of the adults caring for them "to understand, perceive, and respond to [their] bids for assistance and support" (Richter, 2004, p.1). Here, I consider the definition(s), conceptual and theoretical underpinnings, and measurement of sensitive caregiving, before exploring a proportion of the empirical research investigating its influence and malleability.

5.2.2.1 The construct of parental sensitivity

5.2.2.1.1 Defining parental sensitivity

Mary Ainsworth and colleagues conceptualised parental sensitivity as the caregiver's ability to accurately interpret the subtle cues of their child and promptly respond to these signals in an appropriate manner (Ainsworth et al., 1974). Research studies have since characterised this concept as a subtle and intertwined three-part process or chain (e.g., Bornstein & Tamis-LeMonda, 1989; Eshel, Daelmans, Mello, & Martines, 2006):

1. *Observation*: Caregiver observes the child's signals.
2. *Interpretation*: Caregiver accurately interprets the child's signals.
3. *Action*: Caregiver responds promptly and appropriately and the child experiences their needs being met in a predictable manner.

Ainsworth et al.'s (1974) conceptualisation of sensitivity considers broad characteristics of the style of caregiver's response to their child, rather than focusing on a specific checklist of parental behaviours (Mesman & Emmen, 2013). However, in the wider

literature there is subtle variance in the parental behaviours encompassed under different definitions of sensitive caregiving. For instance, more recent depictions of sensitivity promote the importance of parental warmth and positive affect (e.g., Biringen & Easterbrooks, 2012) while such a focus is absent in Ainsworth's definition. Other components included in definitions appear to draw on the writings of Diana Baumrind (Baumrind, 1966, 1991), whose work classified parenting styles based on parental responsiveness and demandingness. Figure 10 offers a depiction of a core set of the parental behaviours considered in the existing conceptualisations of parental sensitivity. Although this diagram illustrates that there is a wide array of parenting behaviours considered under the term 'sensitivity', taken together these behaviours form a style of responsive parenting that is considered to positively promote child development (Landry & Smith, 2011).

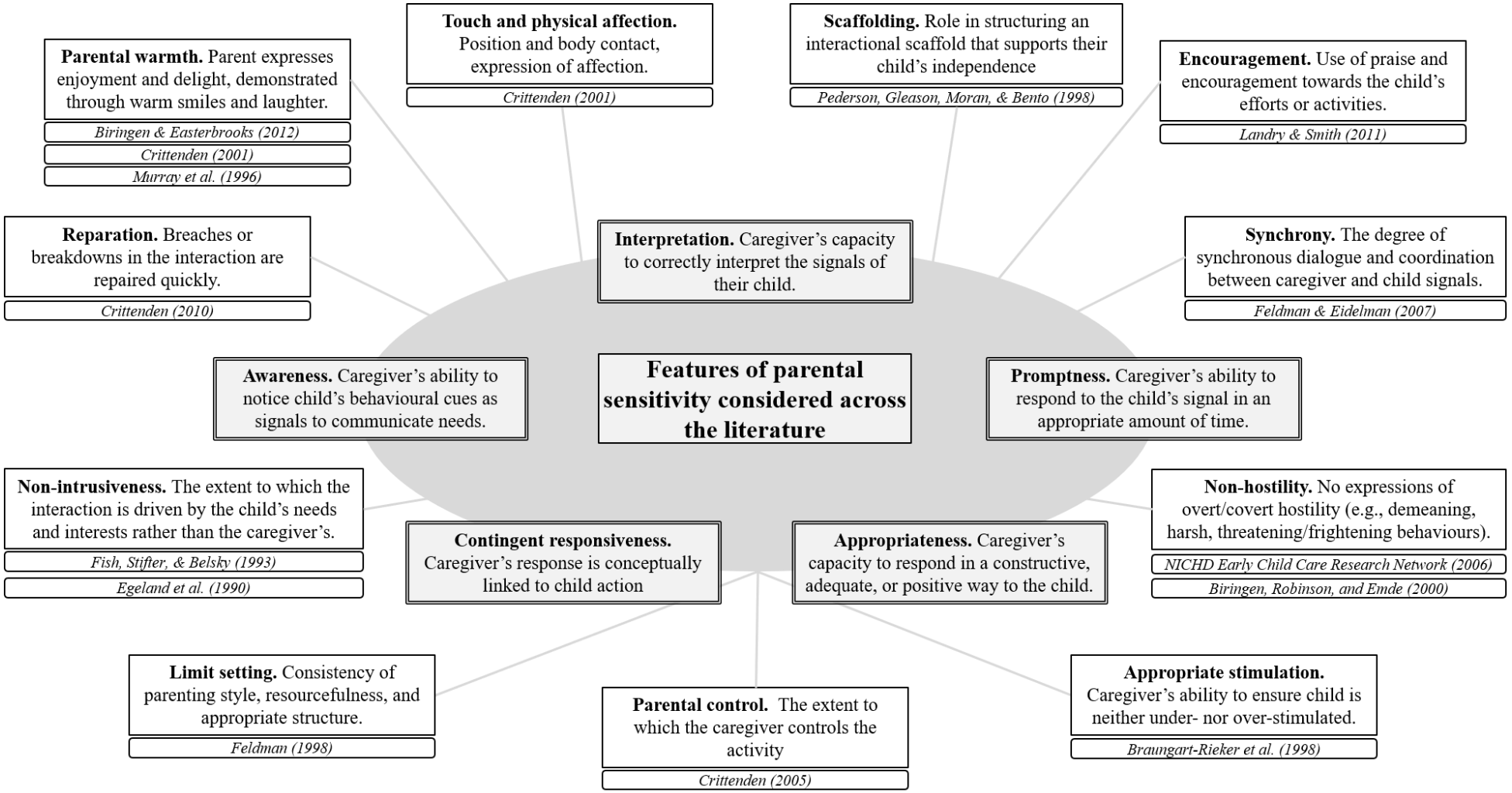


Figure 10. Visual depiction of some of the defining features of parental sensitivity outlined in the literature.

Note. Features of sensitivity that are closely linked to Ainsworth et al.'s (1974) original definition are highlighted in the centre in grey. Parenting behaviours that have been extended upon or newly introduced by other researchers are in white, with references to relevant literature provided where appropriate.

5.2.2.1.2 Theoretical underpinnings of parental sensitivity

Attachment theorists propose that the quality of attachments children form with their closest caregiver(s) can have important implications for their wider development (Ainsworth, Bell, & Stayton, 1971; Ainsworth et al., 1974; Ainsworth et al., 1978; Baumrind, 1967, 1991; Bowlby, 1969). As part of this, sensitive and responsive caregiving has been proposed to promote the internalisation of a secure mental model of primary caregivers, offering children a secure base to explore their environment, trust in and build relationships with others, and engage with learning during future years (van Der Voort, Juffer, & Bakermans-Kranenburg, 2014). The quality of these attachment representations is not considered to be a fixed trait but rather something that can be affected by changes in parenting behaviours (Bakermans-Kranenburg et al., 2003; van Der Voort et al., 2014). It should also be noted that whilst sensitive caregiving is considered to be an important contributor, there are other factors that influence children's development. I focus solely on parental sensitivity in this section as it is a key component of the study at hand.

Foundational studies from Ainsworth and colleagues used intensive, naturalistic home observations (~70 hours of observation per dyad) to reveal very strong links between attachment security and parental sensitivity (r = .78; Ainsworth et al., 1978). However, subsequent meta-analytic evidence reported a much more modest magnitude of association (r = .24; De Wolff & van IJzendoorn, 1997; N = 66 studies). This difference in strength of association could be indicative of several things. First, it could indicate that very intensive methods of observation result in higher correlations between constructs but that such concentrated approaches are less feasible in standard research. Second, it suggests that other aspects of child and family functioning, beyond the attachment framework, may also be associated with parental sensitivity (Moran, Pederson, & Tarabulsy, 2011). For instance, from a socio-cultural perspective, it has been proposed that a parent's capacity to engage in sensitive caregiving is reliant on their capability to operate within their child's zone of proximal development (Vygotsky, 1978, 1987). This is seen as important not only for recognising and responding to their child's focus and interests but also for acting within their child's developmental capacity, to allow the caregiver to modify their behaviours as their child's zone changes (Rodrigues et al., 2021). Overall, through the lens of attachment theory, a caregivers' capacity for sensitive caregiving is just one of several important constructs related to parenting that is thought to contribute to the development of secure attachment relationships.

5.2.2.1.3 Measuring parental sensitivity

There are a variety of approaches available to assess parental behaviours related to sensitive caregiving (Tamis-LeMonda & Baumwell, 2011).

Questionnaires can be especially helpful when considering the effects of parenting in large-scale research, such as cohort studies. However, items often focus on specific parenting behaviours that are assumed to act as an index of parental sensitivity, rather than capturing the quality of parent-child interactions themselves (Tamis-LeMonda & Baumwell, 2011). For example, the Parental Involvement Scale (Sanson et al., 2002) measures the frequency with which parents engage in specific caregiving activities, such as reading, telling a story, and running errands together in a week. The Baby Care Questionnaire (Winstanley & Gattis, 2013) measures parental attunement (e.g., close physical contact) and enforcement of structure (e.g., regularity and routine in infant care) during their child's crying, feeding, and sleeping routines. Such measures can help to shed light on the frequency of specific parenting practices related to supportive parenting. However, they are limited in what they can tell us about the nuances in quality of caregiving. Due to their self-reported nature, they may also be vulnerable to methodological issues related to social desirability bias and individual differences in the interpretation of items.

Where time, training, and funding allows, observational measures may offer a more objective review of parenting. This may be particularly important when assessing sensitivity given that caregivers can only self-report on their responses to the child's signals that they observe, and not those they miss or misinterpret (Rodrigues et al., 2021). Previous research indicates that observational measures of parenting may predict child outcomes more strongly and consistently than self-report measures (Zaslow et al., 2006). Most typically, caregiver-child dyads will be videotaped during an unstructured, semi-structured or structured task (e.g., play time, tidy up, puzzle completion), before the interaction is coded for the characteristics of interest using a pre-defined coding approach.

Seminal work from Ainsworth and colleagues proved formative to the field of parenting and child development, and especially to the ways in which we measure parental sensitivity. Since the development of the Sensitivity-Insensitivity to Infant Signals and Communications scale (Ainsworth et al., 1974), an abundance of different observational measures of sensitivity have been developed. Many of these instruments assess key features of the original definition of sensitivity put forward by Ainsworth but several also introduce

new elements of parental behaviours, as indicators of sensitivity, and vary in the populations and settings they are intended to be applied to (Mesman & Emmen, 2013). Approaches undertaken in observational coding can vary and include macro-level global scoring (e.g., on a five- to ten-point scale), event-based coding (e.g., measuring the frequency or duration of behaviours), and micro-level sequential coding (e.g., measuring the frequency, co-occurrence, and dependency of interaction) (Tamis-LeMonda & Baumwell, 2011). A comprehensive systematic review undertaken by Mesman and Emmen (2013) revealed 50 different measures of parental sensitivity that featured marked differences in the measurement of the construct. Through their searches, they identified a selection of measures that were most well-used in the field. Table 29 provides an overview of those measures that were identified as being suitable for the assessment of caregivers with toddler- and preschool-aged children. For completeness, a description of the Ainsworth scale (most suitable for infants) is also included in the table.

Table 29

An Overview of Widely Used Measures of Parental Sensitivity

Measure	Approach to scoring	Broad description	Key reference
Ainsworth Maternal Sensitivity Scale	One 9-point scale	Awareness of the baby's signals, accurate interpretation of the baby's signals, appropriate responses to signals, and prompt response to signals.	Ainsworth et al. (1974)
The Child-Adult Relationship Experimental Index	A sum of 7 aspects of parent behaviour	Measures parental sensitivity, control, and unresponsiveness through facial expression, vocal expression, position and body contact, expression of affection, pacing of turns, control of the activity, and developmental appropriateness of the activity.	Crittenden (2005)
Coding Interactive Behaviour	A global score derived from 42 5-point scales	The scale scores parent, child, and dyadic behaviour. In total, 21 parent codes are used to score parental sensitivity, intrusiveness, and limit setting.	Feldman (1998)
Erickson scales	Seven 7-point rating scales	Scales include supportive presence, intrusiveness, hostility, sensitivity and timing of instruction, confidence, and clarity of instruction.	Erickson et al. (1985)
Emotional Availability Scales	Six dimensions scored on Likert scales	Scales assess caregiver sensitivity, structuring, non-intrusiveness, and non-hostility. Behaviours of interest include appropriate and positive affect, responsiveness, interference, stimulation, direction, and hostility of voice or facial expression.	Biringen, Robinson, and Emde (2000)
Maternal Behaviour Q-Sort	A global score derived from 90 cards of behaviours	Behaviours are sorted into piles deemed to be highly uncharacteristic to highly characteristic of the parent. Key indicators of sensitivity include reading and response to child signals, synchronicity, scaffolding, dividing attention, and acceptance.	Pederson, Moran, and Bento (1999)

5.2.3 Parental sensitivity in the context of child outcomes and intervention

Building on its theory, defining features, and measurement, I now consider ways in which parental sensitivity has been linked to child outcomes and used as a target for intervention to promote child and family functioning.

5.2.3.1 Parental sensitivity and its role in children's development

Decades of research has provided evidence of the cascading effects of parenting on children's physical, emotional, cognitive, and behavioural development (Britto et al., 2017; Donelan-McCall, 2017; Sroufe et al., 2010). Exposure to a supportive and sensitive parenting environment has not only been identified as a mechanism through which to promote positive child outcomes, but also to buffer against exposure to adverse individual, family, or sociocultural circumstances (Teti et al., 2017). Here, I focus in on research findings that have explored links between parental sensitivity and a) children's attachment security, and b) children's behaviour problems.

5.2.3.1.1 Parental sensitivity and attachment

One of the foundations of attachment theory is that attachment security arises from a caregiving environment that includes accurate, prompt, and consistent parental responses to children's cues and signals. One proposed mechanism for such an association is that children who consistently experience their closest caregiver(s) perceiving and understanding their intentions and distress in a supportive and empathic way are more likely to feel that their needs have been adequately met, leading to the internalisation of a secure mental model of their caregiver(s) (Juffer, Bakermans-Kranenburg, et al., 2017b). Children whose parents struggle to correctly interpret their cues and respond in a prompt and contingent way may be more likely to form an attachment style that is marked by inconsistent or unresponsive caregiving. However, since Bowlby's original proposal that a sensitive caregiving style is a critical antecedent of attachment security (Bowlby, 1969/1982), studies exploring such associations have varied in the ways in which they conceptualise and measure parental sensitivity. Very few studies have used such intensive approaches to observing parental behaviour as Ainsworth and colleagues did. Meta-analyses offer us a systematic approach to studying this large and inconsistent body of research, to better understand whether children who experience greater parental sensitivity are more likely to exhibit a secure attachment style. In their meta-analytic evidence, Goldsmith and Alansky (1987) reported a mean effect size of d = .36, De Wolff and van IJzendoorn (1997) reported a combined effect size of

r(1,097) = .24, and Atkinson et al. (2000) reported an overall effect size of *r*(2,243) = .27 linking caregiver sensitivity and attachment security. Unlike the rather deterministic conclusions of Ainsworth and colleague's studies, these findings indicate that parental sensitivity is an important but not exclusive condition for the development of a secure attachment.

5.2.3.1.2 Parental sensitivity and child behaviour

Empirical research has also established links between early quality of parental caregiving and the long-term development of children's behaviour problems (Fearon et al., 2010; Miner & Clarke-Stewart, 2008). Such associations have been found in both biological and adoptive family relationships, adding weight to the idea that studies are detecting a true effect of the quality of the environment on children's development, rather than genetic influences of biological parents on their offspring (Stams, Juffer, & van IJzendoorn, 2002). Several mechanisms that have been suggested to help explain the relationships between parenting style and children's functioning appear to draw on components of attachment theory. Theorists have suggested that children exposed to conflictual, inconsistent, or unresponsive parenting may 'act out' in an attempt to engage their unresponsive caregiver, to express resentment or distrust of their caregivers, in an effort to maintain distance from others, and/or because they have less to lose in disobeying their caregivers (Donelan-McCall, 2017; Niccols & Feldman, 2006). Scott and Gardner (2015) proposed that such parenting styles may lead to confusion in young children that may prevent them from learning the effects of their behaviour on others. The bidirectional nature of the parent-child relationship must also be kept in mind, with the awareness that more challenging child behaviour may elicit harsher or more critical parenting (Scott & Gardner, 2015).

In considering some of the mechanisms that might underlie the relationships between parental sensitivity and children's behavioural outcomes, and the emphasis placed on attachment theory thus far, a perspective offered by DeKlyen and Greenberg (2016) is helpful to consider. They stress that the "enthusiasm to utilise attachment theory has at times led to a fruitless search for a 'Holy Grail' of psychopathology, seeking main effects of infant attachment on later psychopathology" when instead it generally appears to "exert its influence in the context of other risk factors" (p.640). A relatively recent series of meta-analytic reviews, exploring the predictive significance of early attachment on child development, supports this notion that while early attachment security is important at a population level, it does not necessarily determine later outcomes at an individual level. Estimates of associations

between attachment-related individual differences in early life and externalising symptomatology (d = 0.31, CI 0.23, 0.40; Fearon et al., 2010), internalising symptomatology (d = 0.15, CI 0.05, 0.25; Groh et al., 2012), and social competence (d = 0.39, CI 0.32, 0.47; Groh et al., 2014) were found to be significant but relatively modest in magnitude (using traditional thresholds of effect sizes). Thus, my focus on the relationship between insensitive parenting and children's behavioural outcomes and its possible links to attachment security comes with an awareness of the complex and multifaceted nature of these relationships. Multiple other factors both within and external to the child and family are likely to buffer or potentiate such risks (DeKlyen & Greenberg, 2016), with attachment security interacting with these factors to influence outcome (Sroufe et al., 2010). For instance, other child-level (e.g., gender, age, temperament, emotion regulation), parent-level (e.g., self-confidence, own attachment histories), and social contextual risks (e.g., poverty, social support) will be inflicting their own degree of influence on family processes and outcomes too (e.g., Fearon et al., 2010; Fearon & Belsky, 2011).

5.2.3.2 Promoting parental sensitivity through intervention

Next, I consider the ways in which caregivers can be supported in improving behaviours related to parental sensitivity in the earliest years of their children's lives.

An especially important characteristic of parenting is its malleability. Because of this, caregivers fill a powerful place in children's lives by holding the potential to be the primary agents for change in their children's development. However, promoting good quality early parental care in ways that cater for a diverse array of contextual factors as well as caregivers' own attachment histories, beliefs, behaviour, mental health, and genetic make-up is a significant challenge facing the field (Donelan-McCall, 2017). As was discussed in Chapter 1 of this thesis, parenting programmes with varying theoretical orientations, modes of delivery, and targets for change have been found to have positive effects on family and child outcomes both in the short and long-term (e.g., Bernard et al., 2012; Moss et al., 2011; Smyke, Zeanah, Fox, Nelson, & Guthrie, 2010). However, rather than aiming to develop a one-size-fits-all intervention, there has instead been a push to better understand which interventions work best for whom, when, and why. Continuing to enhance our understanding in this area could enable us to develop a portfolio of programmes that are responsive to the distinct needs of children and families through the family life course. I will now focus on the potential of interventions delivered in the earliest years of children's lives.

5.2.3.2.1 The power and potential of early intervention

Parenting interventions delivered between pregnancy and three years of age are considered to occur during a sensitive window of children's development. A key reason for this is that children undergo rapid neural, physiological, and psychological development between conception and age three years (Hannon, 2003; Winston & Chicot, 2016). The power of environmental inputs during this time of increased neural malleability mean that the effects of children's environmental exposures in their earliest years are considered by some to bear an even greater influence than experiences occurring later in life (DePasquale & Gunnar, 2020). Empirical research indicates that early attachment security interacts with the quality of children's later caregiving experiences in shaping their developmental outcomes, whereby the maintenance of good quality early care can provide a stronger foundation for children's development (Beijersbergen, Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2012; Belsky & Fearon, 2002). Maintaining sensitive caregiving through children's earliest years is regarded as especially important because internal working models are thought to still be rather rudimentary during this life stage (Thompson, 2018). The quality of caregiving during this time is considered to not only have the potential to buffer against stress and adversity but also promote positive child outcomes. Not only this, sensitive caregiving may also have positive economic impacts. A recent longitudinal study revealed that sensitive early parental care predicted significantly lower costs for families and education, health, social, and justice services through to early adolescence (most sensitive quartile mean £1,619, least sensitive quartile mean £21,763; Bachmann et al., 2021). Interventions targeting parental behaviours during children's earliest years are important for families and wider society. In a meta-analysis of 70 studies, short-term, attachment-based interventions were found to have the potential to promote parental sensitivity (d = 0.33; Bakermans-Kranenburg et al., 2003). However, there are relatively few evidence-based early interventions that have been found to scale successfully and lead to positive and enduring effects for families in the UK.

5.2.3.2.2 Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline

Here I provide an overview of one intervention that shows promise in improving family and child functioning. The Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD; Juffer et al., 2008b) is a manualised, interaction-focused intervention that aims to promote two aspects of parenting: parental sensitivity and a sensitive discipline style.

Theoretical underpinnings. The VIPP-SD intervention assumes that the enhancement of parental sensitivity and the promotion of a sensitive discipline style will benefit the parent-child relationship, in turn reducing child behaviour problems (Juffer, Bakermans-Kranenburg, et al., 2017b). It draws on attachment theory by proposing that increased sensitivity from caregivers will promote a secure child-caregiver attachment relationship (Ainsworth et al., 1978; Bowlby, 1969/1982). Components of the intervention also consider behavioural models of learning that relate to social learning and coercion theories (Patterson, 1982). These suggest that family interactions that are characterised by inconsistent or harsh parenting styles may lead to the establishment of 'coercive cycles', whereby caregivers reinforce children's negative behaviours by giving in or not following through, potentially increasing the likelihood of enduring child behaviour problems (Shaw & Bell, 1993). The VIPP-SD intervention instead encourages caregivers to focus on reinforcing children's positive behaviours and promotes a clear, consistent but non-harsh response to children's challenging behaviour.

Key components of intervention delivery. VIPP-SD is comprised of six sessions of video-feedback delivered in the family home. Meta-analytic evidence indicates that short-term interventions (<16 sessions) with a narrow, behavioural focus may be more effective than lengthier programmes (Bakermans-Kranenburg et al., 2003). The programme uses video-recorded interactions of the parent-child dyad during 'typical' family moments (e.g., playing together, reading a book, mealtime). These recordings are then used by the programme therapist to deliver 'feedback' to the parent about their child's behaviour and the parent-child interaction. A stop-start approach is employed when reviewing the videos to allow the therapist to clearly highlight to the caregiver particularly positive and/or subtle moments shared with their child. Undertaking the intervention in the family home is hoped to provide a more representative picture of the family's authentic, day-to-day interactions (Juffer et al., 2008b). Further to this, receiving feedback in a familiar environment, where caregivers and their children likely feel safe and comfortable, may allow caregivers to feel more able to practice new (or extend upon existing) parenting behaviours rather than feeling exposed by the therapist's feedback. Together, the home-based and video-feedback components of the programme's delivery are considered to increase the likelihood of caregivers being able to integrate what they learn during the intervention into their daily interactions with their child (Bakermans-Kranenburg et al., 2003; Juffer et al., 2008b; Juffer, Struis, et al., 2017).

VIPP-SD effectiveness. The VIPP-SD programme's effectiveness has been tested across diverse populations including mothers with postnatal eating disorders (Woolley, Hertzmann, & Stein, 2007), mothers exhibiting an insensitive parenting style (Kalinauskiene et al., 2009), children who have experienced or are at risk of maltreatment (Moss et al., 2011; Pereira, Negrão, Soares, & Mesman, 2014), children at risk for externalising behaviour problems (van Zeijl et al., 2006), children with autism (Green et al., 2015; Poslawsky et al., 2015), adoptive families (Juffer, Bakermans-Kranenburg, & van IJzendoorn, 2005), and in home-based childcare settings (Groeneveld, Vermeer, van IJzendoorn, & Linting, 2011). Juffer, Bakermans-Kranenburg, et al. (2017b) drew together these findings in a meta-analysis involving 1,116 parents and caregivers, reporting a combined effect size of $d = .47$ for improvements to sensitive parenting, $d = .36$ for attachment security, and $d = .26$ for child behaviour problems. These effects appeared to be sustained over a longer-term follow-up, with a reported effect size of $d = .25$ for child outcomes.

These findings provide promising evidence for the effectiveness of the VIPP model in not only improving caregiving behaviours but also aspects of internal and external child functioning. Closer investigations into these apparent shifts in internal functioning could enhance our understanding of the mechanisms underlying such changes. Previous studies that have reported positive effects of VIPP-SD on attachment have assessed the construct at 12-24 months of age using laboratory-based, observational measures designed to activate the child's attachment behavioural system in order to classify their attachment security (Cassibba, Castoro, Costantino, Sette, & van IJzendoorn, 2015; Juffer et al., 2005; Klein Velderman, Bakermans-Kranenburg, Juffer, & van IJzendoorn, 2006; Moss et al., 2011). However, employing other methodologies that aim to access a broader landscape of individuals' internal systems – beyond attachment security – may offer us unique insight into the ways in which environmental enrichment, brought about by effective parenting interventions, can impact upon the ways children understand and respond to their worlds. Theoretically, it is widely regarded that parental sensitivity and attachment relationships are both important influences in the internal systems of children's functioning. However, because these internal systems are not directly observable, we are faced with a challenge in assessing them. One way to try to do this is through the examination of representations, elicited through narratives. The possibility of undertaking such an approach in the context of measuring changes in parental behaviours is considered next.

5.2.4 Measuring intervention effects through children's representations

5.2.4.1 Theoretical underpinnings of internal representations

Under an attachment theory perspective, the quality and content of people's interactions with the external world are considered to influence the formation and modification of their internal cognitive frameworks or so-called 'internal working models' (Bowlby, 1969, 1969/1982; Bretherton & Munholland, 2008). As discussed in the previous chapter, there is variance in the precise ways in which these concepts are characterised. One way to think of internal working models is as a set of expectations, formed out of experiences with different attachment figures that contribute to the anticipation and understanding of future attachment and social relationships (Society for Emotion and Attachment Studies, 2021). Children have the capacity to form multiple attachment relationships with the different caretakers around them. The attachment relationships they form with their closest caregivers are considered to be especially important, given that these key figures represent their most familiar and stable relationships, regardless of the quality of this relationship. Thus, the experiences with these closest caregivers form a significant proportion of the experiences that the child is considered to internalise and generalise to other relationships. These internal systems are considered to guide children as they predict, interpret, and respond to the world around them. It has been proposed that the content and organisation (i.e., coherence) of these internal representations may be contingent on the caregiving environment and interaction patterns children experience with their closest caregivers (Bretherton & Munholland, 2008; Waters, Raby, Ruiz, Martin, & Roisman, 2018). Previous research has indicated that children's capacity to generate coherence and constructive meaning of their world, and their ability to make sense of emotional events, may be associated with wider elements of their development, including their emotion regulation, behavioural development, social interactions, and attachment security (Bretherton & Munholland, 2008; Oppenheim, 2006).

5.2.4.2 Assessing constructs related to internal representations

Gaining a better understanding of internal representations in the context of parenting programmes may be especially insightful, given that several such interventions are underpinned by attachment theory and/or target sensitivity. However, such investigations have rarely attempted to access constructs thought to be related to young participants' internal worlds. One contributing factor to this may be perceived difficulties in measuring such concepts in children still in early childhood.

Well-established measures of attachment security have been developed for specific populations. To assess attachment security in infants and toddlers (generally 12-24 months), the Strange Situation Procedure (SSP; Ainsworth et al., 1978) has been widely used. The mildly stressful situation (separations from and reunions with a caregiver) is designed to activate the child's attachment system in order to classify their attachment security. For adults, the Adult Attachment Interview (AAI; Main et al., 1985) was developed to assess state of mind with regards to attachment. In this assessment, particular emphasis is placed on respondents' speech patterns, and specifically the extent to which they can engage in coherent discourse, while they recount and reflect upon early attachment-related relationships (Waters et al., 2018). The ability to produce a coherent autobiographical narrative has been linked to the quality of individuals' early caregiving, as well as their later romantic and own parent-child relationships (Holland & Roisman, 2010; Shlafer, Raby, Lawler, Hesemeyer, & Roisman, 2015; Steele et al., 2014). The use of these measures has allowed us to gain a better understanding of developmental trajectories in the context of attachment, as well as ways in which attachment representations may be changed by interventions, at least for infants (e.g., Juffer, Bakermans-Kranenburg, et al., 2017b). However, classifying attachment security offers us just one way to understand aspects of individuals' internal systems. Looking at measures of representations more broadly might provide us with unique insight into other aspects related to the ways in which people understand and interpret their worlds. Eliciting narrative representations allows us to look at wider aspects of functioning, particularly for children in early childhood. However, relatively speaking representational measures have rarely been utilised in studies exploring the links between parenting and children's behavioural development, and even less so in intervention studies.

5.2.4.3 Measuring the internal representations of young children

As detailed in previous chapters, story stem completion tasks have been utilised with young children to better understand aspects related to their internal representations. In these methods, children are given dolls and encouraged to enact play narratives in response to (typically emotionally evocative or conflictual) story beginnings. Exploring the ways in which children interpret these scenarios, develop their narrative, and offer a resolution is considered to provide us with insight into their experiences, regulation, and the organisation and content of their scripted inner representations (Yuval-Adler & Oppenheim, 2015).

The development of story stem assessment batteries was largely built on attachment theory. Their use is underpinned by the belief that the availability and responsiveness of key

attachment figures feeds into children's internal representations of their caregiving experiences, as well as their own self-perception and regulation of affect and behaviour (Bretherton, 1985; Bretherton & Munholland, 2008). Aspects of children's internal representations are considered to not only influence the themes they narrate and enact but also the organisation and coherence of their narratives. It has been proposed that children who experience unresponsive or insensitive caregiving may be more likely to have distorted, confused, or biased attachment-related thoughts that produce a disorganised and incoherent narrative (Yuval-Adler & Oppenheim, 2015). There is a small body of empirical support for this proposal. For instance, Bretherton, Prentiss, et al. (1990) and Sher-Censor and Oppenheim (2004) reported associations between children's attachment security in infancy and the coherence of their narratives at age three and four-and-a-half years respectively. However, most studies that have applied story stem methodologies have used concurrent designs with samples of families with no contextual or clinical risk factors (Yuval-Adler & Oppenheim, 2015). Thus, there is a need for studies to employ longitudinal designs, particularly with families considered to benefit from support, to enhance our understanding of how children's early experiences may be reflected in their later narrative representations.

As considered earlier, it may be particularly insightful to include measures of internal functioning in assessments of parenting interventions. Such an approach has only been undertaken in a handful of studies. For instance, Reese et al. (2020; $N = 76$) examined children's narrative responses following a parenting programme that aimed to enhance the quality of parental conversations with their two-year-olds. The programme placed emphasis on parents' use of praise and elaboration, as well as the frequency, sensitivity, and supportiveness of their conversations with their children. At age 11, children's narratives were measured using an interview where they were encouraged to tell a story about the 'best' and 'worst' time in their lives. Children whose parents had participated in the intervention exhibited more coherent and elaborative narratives (ηp^2 .10 - .12; medium-sized effects). Further to this, Olds et al. (2004; total $N = 641$; intervention $n = 197$) used eight story stems from the MacArthur Story Stem Battery to explore the narratives of six-year-olds in a follow-up study of the Family Nurse Partnership programme (Olds, 2006). In low-resourced families, they reported that children whose parents had received the programme told less aggressive ($d = -.25$) and more coherent ($d = -.34$) stories in response to the narrative prompts. These findings indicate that enhancing aspects of parenting through intervention may have knock-on effects on elements of children's internal worlds. However, neither study explored what

factors may have mediated the relationships found (although both speculate it is due to improvements in parenting). Further to this, although both studies aimed to promote aspects of the parent-child relationship, neither intervention primarily targeted constructs closely related to attachment theory (e.g., parental sensitivity). Exploring these relationships may help us to better understand the ways in which sensitivity-focused programmes can promote elements of the cognitive frameworks that are said to guide us through life.

5.2.5 Aims of the current study

The current study aims to draw together key concepts related to sensitive caregiving, children's behavioural development, and the measurement of children's internal representations to help us better understand the influences of early parenting experiences on children's internal and external worlds. This study will first explore associations between children's experiences of early parental sensitivity and constructs derived from their narrative representations two years later. Second, it will examine whether the effects of a brief, video-feedback parenting intervention on children's behavioural development and narrative coherence are mediated by improvements to parental sensitivity post-intervention.

5.3 Method

5.3.1 Study design

The present study used a longitudinal, experimental design. Data for this study were collected as part of the Healthy Start, Happy Start trial (HSHS; O'Farrelly, Barker, et al., 2021; Ramchandani et al., 2017). More detailed descriptions of the trial's design, eligibility criteria, recruitment, sample, and measures are provided in Chapter 2, Section 2.3.

5.3.2 Participants

Participants were a subsample of children and their primary caregivers who were taking part in the HSHS study. The subsample was determined by the availability of observational data that I could code for parental sensitivity. Selection was based on participating families who had had their baseline assessment conducted by a research assistant other than me in an effort to reduce researcher bias. It should be noted though that whilst significant efforts were made to reduce bias, I had met 91 (49%) families included in analysis at one or both of their follow-up assessments thus not eliminating bias entirely. Participating families were also required to have fully completed all three research assessment visits (at baseline, 5-month, and 24-month follow-up) for inclusion in the present analysis.

Screening and recruitment of the full HSHS sample was conducted across six NHS sites in the South and South East of the UK. Through health and community services, caregivers with children aged 12 to 36 months were asked to complete a screening questionnaire (SDQ; Goodman, 1997) that measured children's behaviour problems. Families whose children scored in the top 20% of population norms for externalising behaviours were contacted to see if they were willing and eligible to participate in the main randomised controlled trial. Following their baseline assessment and enrolment into the study, families were randomised into either the treatment (VIPP-SD plus usual care) or control (usual care alone) arm of the trial. Families were followed up twice after their baseline visit, at five-months and 24-months post-randomisation.

5.3.3 Measures

This section provides detail on the measurement of parental sensitivity, parent-reported child behaviour, and self-reported child outcomes. For clarity, Figure 11 provides a visual representation of the timing of data collection across the three time points. Both the measures of parent- and child-reported behaviour have been introduced in earlier sections of this thesis

(for instance, see Chapter 2, Section 2.3.8 and Chapter 3, Section 3.3.3.1). However, parental sensitivity has not yet been included in analyses. Given this, providing a detailed description of the measurement of parental sensitivity forms the primary focus of this section, with briefer detail provided for the other measures of interest.

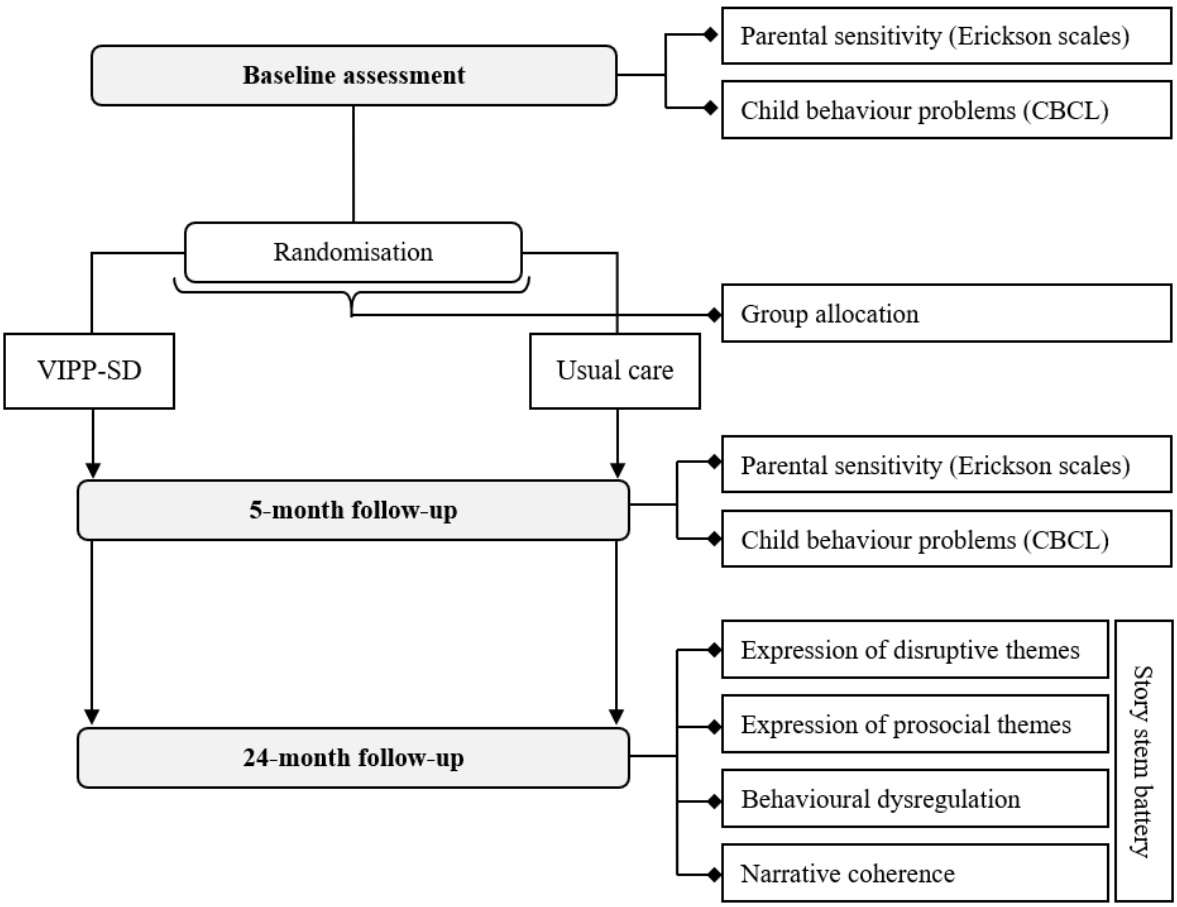


Figure 11. Timing of data collection to explore links between early parental sensitivity and children's later behavioural and representational outcomes.

5.3.3.1 Parental sensitivity: The Erickson scales

Parental sensitivity was assessed through the observational coding of video-recorded parent-child interactions. Behavioural coding can offer rich data and unique insight into family behaviours compared to self-reported methods, which can be susceptible to social desirability bias (Zaslow et al., 2006). Observational approaches also have the potential to capture nuanced information that participants are unaware of or that researchers would not have the foresight to ask about (Pesch & Lumeng, 2017). In this section, I first describe the process of obtaining the parent-child interactions and then detail the procedure undertaken to apply the coding scales to the data.

5.3.3.1.1 Capturing parent-child interactions

Parental sensitivity was assessed using a free-play parent-child interaction, undertaken by families as a part of a "three bag task". The task was approximately 15 minutes in duration. Participating families were provided with a different bag of age-appropriate toys for each of the three interactions and were asked to only play with the provided toys. For the first bag the child was recorded playing alone, for the second the parent and child played together, and for the final bag the parent was asked to engage in a discipline interaction with the child whereby they were instructed to prevent the child from touching any toys. To answer the research questions of the current study, within the boundaries of resource provided by a single PhD student, only the data captured through the 'playing together' interaction was used in analysis. This was a relatively unstructured, low-stress interaction where the caregiver was encouraged to play with the bag of toys "as they normally would" with their child. Compared to the other two available video recordings, this task elicited close parent-child communication. This style of parent-child interaction is in keeping with the kinds of day-to-day interactions commonly used in the coding of parental sensitivity. The bag of toys contained an interactive picnic basket, a car, and two hand puppets. The caregiver and their child were video-recorded interacting for five minutes.

5.3.3.1.2 Coding of parent-child interactions: Scales, training, procedure, reliability, video selection.

5.3.3.1.2.1 Coding scales

In this study, two core components of primary caregivers' sensitive responsiveness were coded during the 'playing together' parent-child interaction: their supportive presence and their intrusiveness. These behaviours were assessed using the revised Erickson scales (Egeland, Erickson, Clemenhagen-Moon, Hiester, & Korfmacher, 1990; Erickson et al., 1985). Previous studies that have used these scales report acceptable psychometric properties of the measure including good to excellent inter-rater reliability (Cents et al., 2014) and acceptable test-retest reliability over a period of nine to twelve months (*r*s .56-.84; Stams, Juffer, Rispens, & Hoksbergen, 2000).

Supportive presence refers to the caregiver's availability to their child, their expression of emotional support, acknowledgement and extension of their accomplishments, and the extent to which they offer encouragement, confidence, enthusiasm, and reassurance, with a positive regard and calming presence. Parental supportiveness was coded on a seven-

point scale, ranging from 1 (*completely fails to be supportive*) to 7 (*skilfully provides support*). *Intrusiveness* relates to the extent to which caregivers display disrespect for their child's autonomy, characterised by caregivers' interference with their child's needs, desires, interests, or behaviours. Parental intrusiveness was also scored using a seven-point scale, ranging from 1 (*nonintrusive*) to 7 (*highly intrusive*). The Erickson scales provide behavioural descriptions for each score on the scale to guide scoring.

For this analysis, an overall 'parental sensitivity' score was generated using the Erickson scales of supportiveness and intrusiveness. In keeping with the approach taken by other research teams, including Lucassen et al. (2015), Kok et al. (2015), Mileva-Seitz et al. (2015), and Windhorst et al. (2015), sensitivity scores were generated using the following process: 1) the intrusiveness scale was reversed so that higher scores indicated a more cooperative parenting style, 2) both the supportiveness and intrusiveness variables were standardised, 3) the standardised scores were averaged (summed and divided by two) to form an overall parental sensitivity score.

5.3.3.1.2.2 Training phase

To complete the observational coding, I undertook an intensive face-to-face training course, conducted by Professor Marian Bakermans-Kranenburg, an expert in the field of parental sensitivity. The initial training was conducted across five days and involved being trained on both the Ainsworth scales of sensitivity and cooperation (Ainsworth et al., 1974) and the revised Erickson scales of supportiveness and intrusiveness (Egeland et al., 1990; Erickson et al., 1985). In the first half of each day of training, we watched videos of caregiver-child interactions together and used a stop-start approach to consider how to apply the scales to the data, under the guidance of Prof. Bakermans-Kranenburg. The second half of each day was spent coding a set of videos independently. We then reconvened as a group to discuss our scoring, thoughts, and to consider any discrepancies or to seek clarity. This approach continued for the five days of training.

The Ainsworth scales are considered suitable for assessing parental sensitivity when children are in infancy (under 18 months or so), whereas the Erickson scales are considered suitable for assessment in the toddler and preschool years (Mesman & Emmen, 2013). In the current study, 70% of children were aged over 18 months at baseline. Thus, following consultation with Prof. Bakermans-Kranenburg, the decision was made to apply the Erickson scales to the data rather than the Ainsworth scales. However, the constructs underlying the

Ainsworth and Erickson scales and – perhaps more notably – the observational coding skills required to apply them are very similar. The initial intensive training in the Ainsworth scales was considered to provide an essential foundation of skills to then make the transition to training on the Erickson scales for use with older children.

5.3.3.1.2.3 Reliability phase

Assessment of inter-rater reliability was undertaken in two stages. First, reliability on the Ainsworth scales was established using a 16-case reliability set. Then, another 16-case reliability set was used to assess reliability on the Erickson scales. Intra-class correlation coefficients (ICC; Koo & Li, 2016) were run to assess the degree of correlation and agreement between myself and Prof. Bakermans-Kranenburg for both scales (see Table 30). Using widely-accepted thresholds of agreement, reliability scores for this study can be interpreted as fair to excellent (Portney, 2020).

Table 30

Intra-Class Correlation Coefficients for the Ainsworth and Erickson Sensitivity Scales

Scale	Ainsworth scales	Erickson scales
Sensitivity/supportiveness	.88	.70
Cooperation/intrusiveness	.91	.82

5.3.3.1.2.4 Coding procedure

The approach to using the Erickson scales involved noting down commentary and assessment of parent and child behaviours approximately every 5 to 20 seconds of interaction. Coding notes focused on the child's signals and an assessment of the caregiver's identification of the child's cues as well as the promptness and adequacy of their response. Assessment also focused on times the caregiver intruded upon or interfered with the child's exploration and play. Such behaviours included times the caregiver led on the play, did not allow the child time or space to explore at their own pace, and/or unnecessarily redirected the child's attention elsewhere either verbally or physically. The codes were then assessed as a whole and a short summary of both the caregiver's supportiveness and intrusiveness were written. A score was then assigned for each scale, using the descriptions accompanying the scores on the Erickson scales for guidance.

Figure 12 provides an excerpt of data I analysed to give an idea of the kinds of behaviours noted down during coding. Please note, this is a summarised version of the full coding document. The excerpt contains a few comments per minute of interaction, compared to the complete coding sheet for this family which contained 40 separate comments. Some details of the interaction have been altered to prevent participant identification.

Time	What happens?	Notes
00:24	C holding fork and focuses on this, M turns away and tips out contents of picnic basket	M focused on her own play here, misses C's focus on fork
00:37	M picks up food and says 'where does that one go?', C says 'this one's for you' and hands M fork, M asks 'that one's for me?'	M quick to shift to C's play after initially providing direction, responsive to vocalisations
00:56	C says 'apple', M corrects this, C then says 'orange' and M says 'good boy'	Praise offered, correction delivered in a sensitive way
01:02	C places watermelon in plate and M says 'do you know what that one is? It's watermelon'	Extension of C's exploration, responsive to C's actions
01:10	C is holding fork, M says 'and you've got a fork', picks up her own fork and pretends to eat with it	Sensitive extension of C's play, M joining in C's play
01:20	C pretends to eat, saying 'let's eat, apple'. M turns to turn on picnic basket instead	M misses opportunity to join in C's picnic, unnecessarily introduces new kind of play
01:33	C becomes interested in picnic basket, M moves it closer to him so he can get a better look	M facilitating C's exploration here
01:59	The picnic basket says 'can you find another circle?', M asks C if he can find a circle, C says 'no' and M repeats this back to him	M is instructing child a little here, could have seen where he would take play, but then responsive to C's vocalisations
02:03	C picks up two shapes and M names these as he picks them up 'triangle and square'	Positive commentary of C's own exploration, sensitive
02:08	C focused on playing with food, M picks up cups and says 'do you want a drink?'	M shifting direction of play a little here, could have seen where child would take play next
02:22	C picks up cups and says 'cup!', M asks C to find all the square food	Mismatch in interest here, unsupportive of C's current interest and directive
02:28	C puts orange in M's cup, says 'orange' and M says 'ah you made me an orange juice'	M quick to switch program of play to suit C's intentions, supportive
02:42	C is playing with picnic food, M tries to extend play, C reacts negatively to this by screaming, M alters her behaviour and puts the sandwich back	M attempting to extend play here but is then responsive to C's signals when C protests
03:05	C tips food into M's cup, M says 'oh, I've got two'	M joining in C's play here, nice commentary
03:18	C puts food in cup. M says 'does it fit in there? Maybe you could put it in here' pointing to plates	M offers commentary, maybe directs a little unnecessarily, but still responsive to C's actions
03:30	C manages to fit piece in correct slot, M asks if she can have it, C says 'no', so M removes another piece from C's plate, C protests	M misses opportunity to offer C praise here and also moves at a fast pace so that C becomes frustrated, a little interfering
04:01	M a little in 'teaching mode', asking the child to find every thing on the plates	M is engaged with child, but instructing a little rather than matching C's pace
04:14	C holding cup and fork, M turns on picnic basket, asks C to find something orange, C points to fork and M says 'a fork? That's red!'	M leading on play a bit and instructing child, but responsive to C's vocalisations of 'fork'
04:24	C tries to push button of picnic basket again but pushes wrong button. M says 'it's this one', C presses this and produces music	M supportive, providing helpful verbal instructions while still allowing C to try it for himself
04:30	Picnic basket says 'find the squares', M repeats this twice. C points to them and M says 'good boy'	Praise offered but M directing course of play by asking child to listen to picnic instructions
04:39	C picks up picnic piece, M says 'shall we put them in the basket? Let's find the squares'	M is engaged in what C is doing and responds to all vocalisations but still directing play a little
04:49	C continues putting pieces in basket. M says 'are we tidying up now? Have you finished?'	Sensitive, commentary, and now responsive to C taking lead on play

Figure 12. Excerpt of coding notes taken from a coding case completed in the current study.

5.3.3.1.2.5 Allocation of videos for coding

One of the most significant challenges of observational coding is the time, expense, and skill necessary to reach adequate levels of reliability on the selected coding scales. For the current study, the original intention was for coding to be undertaken by multiple PhD

students. However, due to time constraints and challenges in reaching acceptable levels of reliability, I was the only researcher available to undertake coding of this data. This brings significant limitations to the observational coding which will be discussed further in the discussion section of this chapter.

5.3.3.2 Children's representations: The doll's house play story stem assessment battery

The doll's house play story stem assessment battery aimed to elicit children's representations and reports of their own behaviour. For brevity, I have provided an abridged outline of the development, administration, and coding procedures undertaken for this measure here (see Chapter 3, Section 3.3.3.1 for more detailed information). Development of this measure was informed by the design and procedure of the MacArthur Story Stem Battery (MSSB; Bretherton, Oppenheim, et al., 1990) and the Doll's House Play approach (Murray et al., 1999; Ramchandani et al., 2011).

I led on the piloting of this measure with 20 three- and four-year-olds in nursery and home settings. Iterative changes were made to the measure's administration procedure following piloting with each child and discussions within the research group until the final procedure was reached. Three story stems were selected for use with the HSHS sample. Alongside two other researchers, I administered this measure as part of the wider 24-month follow-up assessment. During administration, the researcher sat with the child at a table and explained that they were going to tell three stories together, with the researcher telling the beginning of the story and the child able to tell the ending. If the child was happy to complete the measure, they were presented with 15 dolls and asked to pick the dolls that they thought were most like them and the family members who lived with them. Following a warm-up 'party' story where the core concepts of the measure were introduced to the child (e.g., manipulating the dolls, speaking for the dolls, providing narration, and talking to the researcher), the child was invited to play through the three scenarios (spilt juice, hurt knee, three's a crowd). In each of these scenarios, the doll representing the child was placed in a situation of distress or conflict. The child was then encouraged to finish the story through play and narration with the prompt "show me and tell me what happens now" from the researcher. See Appendices E and F for the complete administration manual and coding scheme.

Coding of the story stem data was completed by myself and three graduate coders. All coders reached a moderate to excellent level of reliability (ICC: .68-.98) on the constructs coded. The coding scales applied to the story stem videos were informed by coding

approaches undertaken in other studies that have used children as informants. The scoring used in the current study was divided into four core sections:

1. *Disruptive content themes:* Aggressive, destructive, and ostracising behaviours.
2. *Prosocial content themes:* Helpful, caring behaviours.
3. *Narrative coherence:* How organised and coherent the child's storytelling was.
4. *Observed behavioural dysregulation:* Ability to regulate behaviour during measure administration (e.g., staying in seat, maintaining concentration).

5.3.3.3 Behaviour problems: Child Behaviour Checklist

The Child Behaviour Checklist (CBCL/1.5-5; Achenbach & Rescorla, 2000, 2001) is a 100-item measure of behavioural difficulties. Primary caregivers were asked to rate how true each described behaviour was of their child in the last two months using a three-point scale (0 = not true, 1 = somewhat true, or 2 = very true or often true). Scoring generated a total score as well as an internalising score (made up of items from the anxious/depressed, emotionally reactive, somatic complaints, and withdrawn syndrome subscales) and externalising score (produced by the attention problems and aggressive behaviour subscales). The CBCL is widely used in research and has been found to be reliable and valid in children as young as one year of age (van Zeijl et al., 2006). Detailed information about the measure's reliability and validity is provided in Chapter 2, Section 2.3.8.

5.3.3.4 Demographic information

Demographic information for participating caregivers and their children was collected through an interview with children's caregivers at their baseline assessment. Variables collected included caregiver age, sex, ethnicity, educational attainment, employment status, and relationship status, and child age, sex, and ethnicity.

5.3.4 Study procedure

Study assessments were conducted in families' homes and generally took between two and two-and-a-half hours to complete. The child's primary caregiver provided consent for their own participation and on behalf of their child. At baseline visits, the caregiver provided informed consent and demographic information before completing two researcher-led interviews. Following this, the caregiver and child played with three bags of toys while being video recorded. These formed three parent-child interactions, of which the 'playing together' interaction was analysed in the current study. Finally, the caregiver completed a battery of

questionnaires (including the CBCL). This procedure was repeated for the 5-month follow-up. At the 24-month follow-up, the procedure was very similar but while the caregiver was completing the questionnaires, the researcher administered a set of direct assessments of the child, including the story stem assessment battery used in the present study.

5.3.5 Intervention

Participating families were randomised into one of two trial arms following their baseline assessment. Brief detail about the randomisation procedure and both groups are outlined here.

5.3.5.1 Randomisation

Following their baseline assessment and enrolment into the trial, participants were randomly allocated (1:1) to either the intervention arm (VIPP-SD plus usual care) or the control arm (usual care) of the trial. The randomisation list was generated by an independent statistician and used block sizes of 2, 4, and 6 varying at random. Randomisation was stratified by recruitment site and number of participating caregivers (one versus two). Participating families were allocated to the next available treatment code. Access to the allocation sequence was restricted and not accessible to any members of the research team.

5.3.5.2 Intervention arm: VIPP-SD

Families who were randomised to the VIPP-SD group were offered six home-based sessions of the Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline programme (Juffer et al., 2008b; Juffer, Bakermans-Kranenburg, et al., 2017b). VIPP-SD is a brief, manualised, home-based intervention that aims to promote parental sensitivity (i.e., their capacity to identify their child's attachment cues and exploratory behaviour and to respond to them appropriately) and a sensitive discipline style (i.e., a consistent but non-harsh response to challenging behaviour). Programme visits were scheduled on a fortnightly basis and were generally one to two hours in duration. The therapists delivering the programme were from predominantly health visiting, nursing, and psychology backgrounds. The programme visits were split into two parts. The first part of the visit involved the caregiver and child being filmed together during interactions that were designed to mimic everyday moments of family life. For example, caregivers were asked to play with toys, read a book, play games, and have a meal with their child. Following the session, the video clips obtained through these interactions were then used by the programme therapist to write a 'script'. This script was used in the following session to help guide

feedback delivered by the therapist when pausing the video interaction every 10 to 30 seconds. The focus of the feedback was informed by the sensitivity and disciple themes outlined in the VIPP-SD intervention manual but was personalised to each family as it was based on the unique interaction happening between that caregiver and child. Fidelity was assessed by rating a random selection of audio-recorded visits for adherence to the manual. Fidelity checks indicated that 94% of visits met the pre-specified threshold for adherence to the manual (reported in O'Farrelly, Barker, et al., 2021).

5.3.5.3 Usual care

Participants in both the VIPP-SD and usual care arms continued to access usual care. In most cases, this care was minimal as there are no standard care pathways through the NHS for early onset behaviour problems. General support available to families was mostly through health visiting and general practitioner services.

5.3.5.4 Blinding

Research assistants were 'blind' throughout data collection and analysis. Participating families were regularly reminded of the importance of the research team not knowing their treatment allocation. In the rare instances where unblinding did occur (2% and 4% of assessments at each follow-up point), recordings of researcher-rated measures were double scored by a second, blinded assessor.

5.3.6 Research aims and planned statistical analysis

5.3.6.1 Research aims

This study built on the findings of the previous two chapters, by extending the focus on parenting to also include a measure of parental sensitivity. The story stem measure of children's narrative representations was used as a complement to parent-reported data of child behaviour (CBCL). Particularly close attention was paid to the narrative coherence of children's stories in analysis given its close theoretical ties to parental sensitivity and attachment – key foundations of the VIPP-SD intervention. Child behaviour was measured at 5-month follow-up, rather than 24-month follow-up for two key reasons. First, from a theoretical perspective, exploring child outcomes at both 5-month and 24-month follow-up allows for an interesting look at immediate and enduring effects of this intervention. Second, practically, the complete data was available for the 5-month follow-up first.

This study had three central aims:

1. To explore associations between early parental sensitivity and children's later narratives in relation to the four scales described in section 5.3.3.2.
2. To explore whether intervention effects on narrative coherence at two-years post-randomisation were mediated by parental sensitivity.
3. To explore whether the intervention reduced child behaviour problems at post-treatment and whether these effects were mediated by parental sensitivity.

5.3.6.2 Hypotheses

Three main hypotheses were proposed for the current analysis:

1. Parental sensitivity at baseline would be associated with more coherent and prosocial stories with fewer disruptive themes and lower behavioural dysregulation in children's responses to the story stem task at the 24-month follow-up.
2. Parental sensitivity measured at the five-month follow-up would mediate the effects (reported in the previous chapter) of VIPP-SD seen on children's narrative coherence at 24-month follow-up.
3. a) Delivery of VIPP-SD would reduce children's behaviour problems measured using a parent-reported questionnaire and b) parental sensitivity measured at the five-month follow-up would mediate these effects.

5.3.6.3 Statistical analysis plan

Statistical analysis was split into five parts.

1. **Power analysis:** First, post-hoc power analyses was conducted using the G*Power programme.
2. **Exploration of data:** Second, the flow and retention of participants was explored before the demographic characteristics of the sample and descriptive statistics of study variables were presented.
3. ***Aim 1: Exploring associations between early parental sensitivity and children's later narrative representations*:** For aim 1, multivariate assumptions of normality were tested before multiple regression models were constructed to explore the associations between early parental sensitivity and outcomes measured through the story stem assessment.
4. ***Aim 2: Exploring intervention effects on narrative coherence (story stems) mediated by parental sensitivity:*** For aim 2, multivariate assumptions of normality were tested

before multiple regression models were constructed to explore associations between group allocation and parental sensitivity and group allocation and children's narrative coherence for this subsample. Further analyses were conducted to explore the two subscales of the sensitivity scale (supportiveness and intrusiveness) separately. Then group differences on the mediator and outcome variables were examined using t-tests before mediation analysis was conducted using the *PROCESS* v3.5 extension (Hayes, 2018). Mediation analysis tested whether VIPP-SD had an effect on parental sensitivity at five-month follow-up (the mediator), and whether any potential changes in the mediator were associated with changes in children's narrative coherence.

5. ***Aim 3: Exploring intervention effects on children's behaviour (CBCL) mediated by parental sensitivity:*** For aim 3, multivariate assumptions of normality were tested before multiple regression models were constructed to explore the associations between group allocation and parental sensitivity and another to explore the associations between group allocation and children's behavioural outcomes. Further analyses were conducted to explore the two subscales of the sensitivity scale (supportiveness and intrusiveness) separately. Then group differences on the mediator and outcome variables were examined before mediation analysis was conducted using the *PROCESS* v3.5 extension (Hayes, 2018). Mediation analysis tested whether parental sensitivity measured post-intervention mediated any relationship between group allocation and children's behaviour at the five-month follow-up.

5.3.6.3.1 Procedure undertaken for mediation analysis

Hayes' (2018) *PROCESS* tool uses ordinary least squares (OLS) regression for model estimation to estimate the indirect effect and its confidence intervals. Mediation models were run even when there was no evidence of an association between X and Y, as correlations between these variables are no longer viewed as a precondition to mediation (Bollen, 1989). All presented regression coefficients were unstandardised as standardised regression coefficients are not recommended for dichotomous antecedent variables (Hayes, 2018). Bootstrap confidence intervals were used to guide inferences about indirect effects, rather than the Sobel test (Sobel, 1982) due to the unrealistic assumptions related to normality and sample size of this test (Field, 2013). This approach to mediation provided an estimate of the indirect effect and its confidence intervals, as opposed to the more traditional Baron and Kenny (1986) approach that mostly focuses on assessing whether a change in *p*-values is

statistically significant (Field, 2013). See Figures 13 and 14 for diagrams illustrating the simple mediation models described in steps four and five respectively.

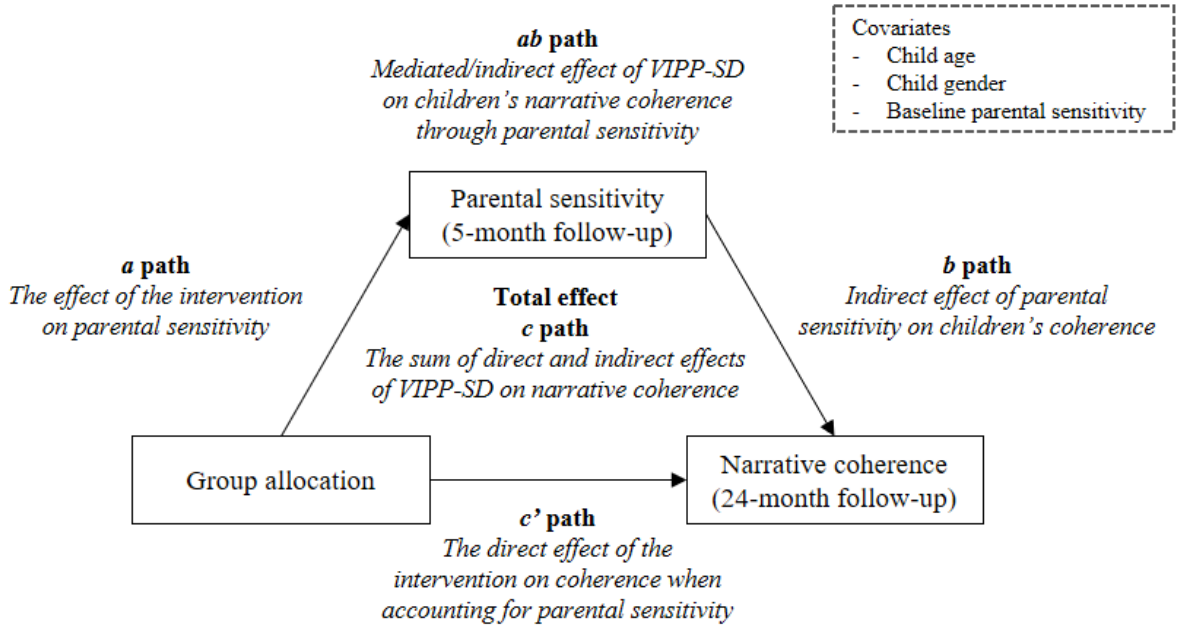


Figure 13. Simple mediation model: The mediating effect of post-intervention parental sensitivity on children's narrative coherence.

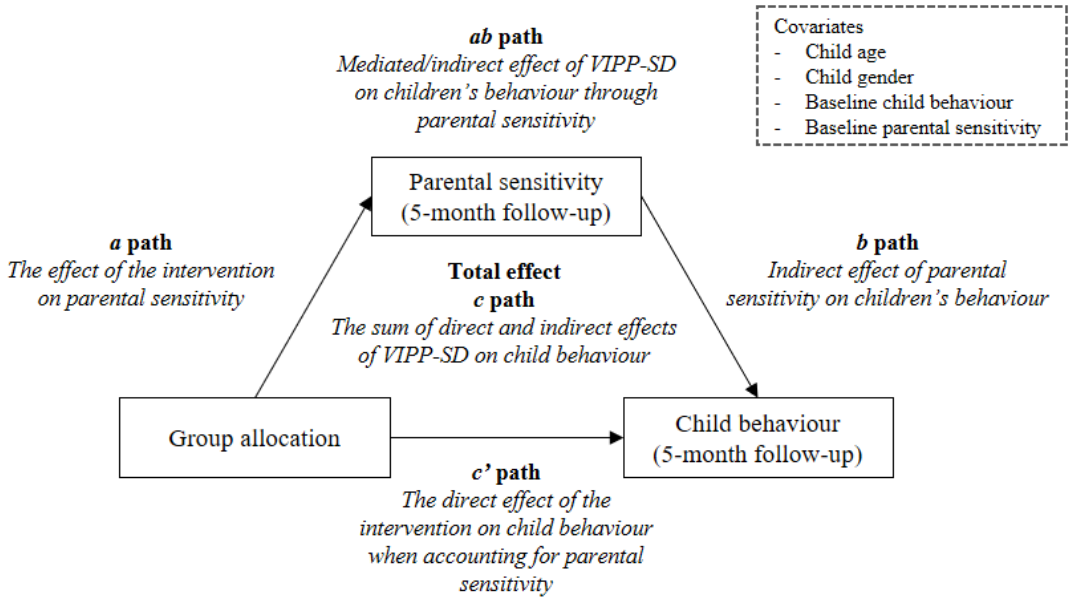


Figure 14. Simple mediation model: The mediating effect of post-intervention parental sensitivity on children's behaviour problems.

5.4 Results

5.4.1 Data included in current analysis

The Healthy Start, Happy Start trial (HSHS) recruited to target (N = 300). Across analyses conducted as part of the current study, smaller subsamples of families were included (ns 154-185). The subsamples were partly determined by overall retention of participants but also by which variables were available (e.g., coded or consented for) in the dataset. Figure 15 displays the flow of participant recruitment and retention, as well as reasons for exclusion from the current analysis.

5.4.2 Power analysis

Estimating the statistical power that a study's sample provides is important to understand the probability of making a Type II error (missing an effect in a population that genuinely exists). The sample size of 300 participants for the full HSHS sample was selected to provide 80% and 90% power to detect standardised effect sizes of 0.36 and 0.42 respectively, at a 5% level of statistical significance, assuming a 20% attrition rate. As the current sample was significantly smaller than this (ns 154-185 families; 51-62% of full sample), a new power analysis was run to estimate the power provided by this reduced sample size. For this, the G*Power tool (Faul et al., 2009) was used. For the multiple regression analyses, with a maximum of five predictors and a sample of 154 families, there was 97% power to detect a medium-sized effect of $f^2 = 0.15$ ($\alpha = 0.05$). Independent sample t-tests were used to explore group differences in narrative coherence (story stems), child behaviour (CBCL), and parental sensitivity (Erickson scales). Samples sizes and estimated effect sizes varied across this set of analyses. For child behaviour, with the available sample of 178 families, there was 41% power ($1 - \beta = 0.41$; $\alpha = 0.05$; two-tailed) to detect an estimated effect size of $d = 0.26$ (informed by the effect size for child behaviour detected in prior meta-analytic evidence of the VIPP intervention). For narrative coherence, with the available sample of 154 families, there was 67% power ($1 - \beta = 0.67$; $\alpha = 0.05$; two-tailed) to detect an estimated effect size of $d = 0.39$ (informed by the effect size found for narrative coherence detected in the larger sample of the previous chapter). For parental sensitivity, with the available sample of 154 families, there was 82% power ($1 - \beta = 0.82$; $\alpha = 0.05$; two-tailed) to detect an estimated effect size of $d = 0.47$ (informed by the effect size for parental sensitivity detected in prior meta-analytic evidence of the VIPP intervention. To detect an effect size of $d = 0.26$ at 80% power for children's behaviour problems, calculations suggest 468 participants would have

been required. To detect an effect of $d = 0.39$ at 80% power for children's narrative coherence, calculations suggest 210 participants would have been required. Therefore, this study is slightly underpowered to detect differences in effects of this magnitude for narrative coherence and significantly under-powered to detect effects on children's behaviour problems.

5.4.3 Descriptive analysis of the study sample and study variables

5.4.3.1 Descriptive analysis of participant recruitment pathways and demographics

Detailed information on the recruitment of participating families is available in Chapter 2, Section 2.3.8. Families were recruited through health visiting, primary care, and community services across six NHS sites. Most participants were recruited through health visiting services (53% total; 37% at developmental reviews and 16% through service mailshots) and children's centres (34%) (see Table 31). The majority of families were recruited through NHS sites in London (78%) with smaller numbers from Oxfordshire, Hertfordshire, and Peterborough (see Table 32).

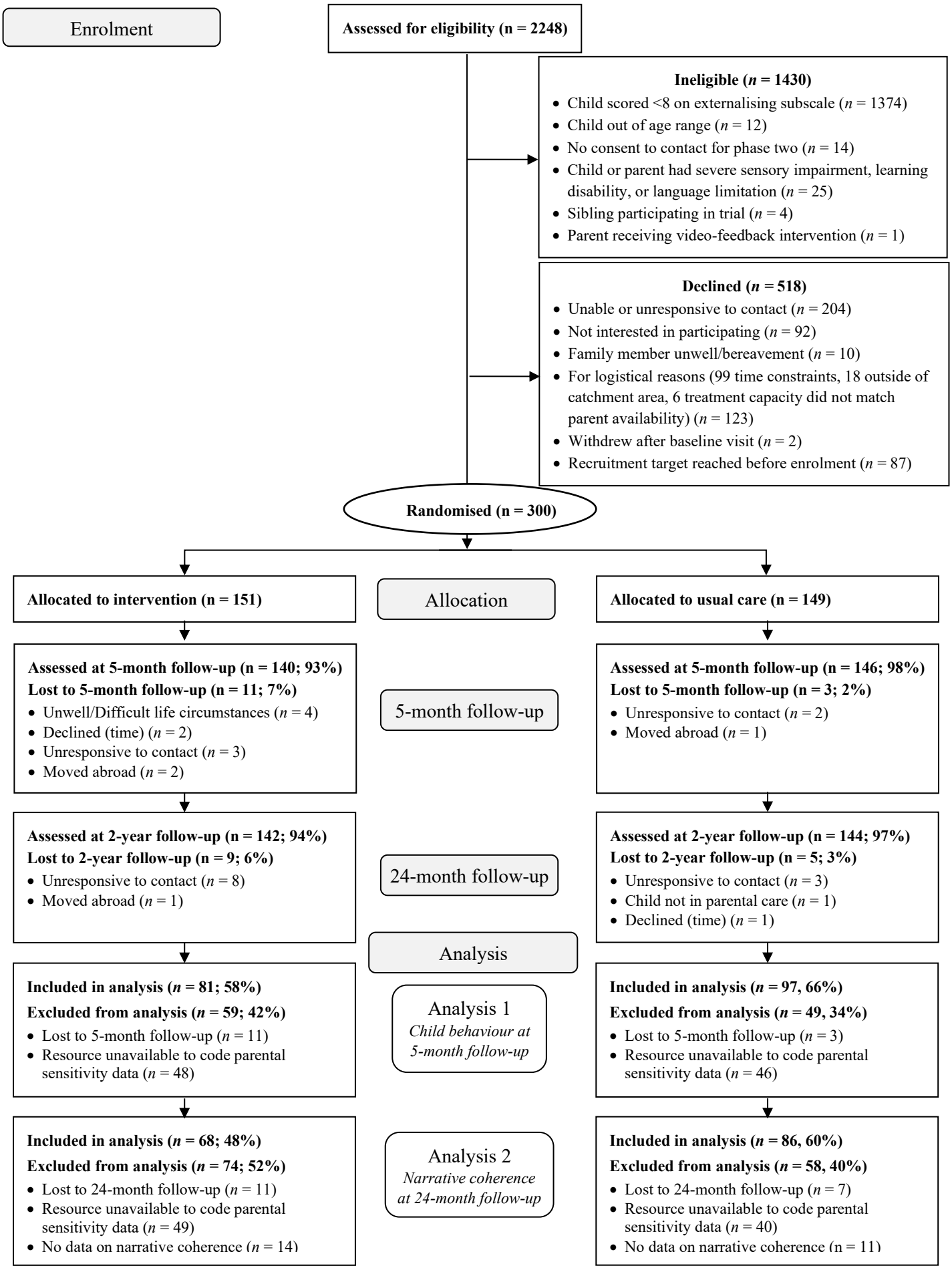


Figure 15. The CONSORT flow diagram for participant data included in the current analyses.

Table 31

Recruitment Pathway of Participants Disaggregated by Treatment Arm

Recruitment route at screening	VIPP-SD (n = 87)	UC (n = 98)	All (N = 185)
Developmental health visitor review, *n* (%)	31 (36)	37 (38)	68 (37)
Health visiting mailshot, *n* (%)	14 (16)	16 (16)	30 (16)
Children's centre, *n* (%)	30 (35)	32 (33)	62 (33)
Other clinic/community venue, *n* (%)	7 (8)	8 (8)	15 (8)
Online advert, *n* (%)	2 (2)	1 (1)	3 (2)
Word of mouth, *n* (%)	1 (1)	2 (2)	3 (2)
Other, *n* (%)	2 (2)	2 (2)	4 (2)

Note. VIPP-SD = Video-feedback Intervention to promote Positive Parenting; UC = usual care.

Table 32

Recruitment Location of Participating Families Disaggregated by Treatment Arm

Recruitment site	VIPP-SD (n = 87) *n* (%)	UC (n = 98) *n* (%)	Total (N = 185) *n* (%)
Barking and Dagenham	1 (1)	3 (3)	4 (2)
Camden	33 (38)	30 (31)	63 (34)
Hertfordshire	2 (2)	2 (2)	4 (2)
Hillingdon	5 (6)	11 (11)	16 (9)
Islington	29 (33)	28 (29)	57 (31)
Oxford	11 (13)	16 (16)	27 (15)
Peterborough	6 (7)	8 (8)	14 (7)

Note. VIPP-SD = Video-feedback Intervention to promote Positive Parenting; UC = usual care.

5.4.3.2 Descriptive analysis of the study sample

Table 33 presents sociodemographic information related to the participating children and primary caregivers for the full HSHS sample (N = 300) and the subsample included in the current study (n = 185).

Children had a mean age of 23.2 months (~2 years; SD = 6.8 months) at baseline, 29.1 months at 5-month follow-up (~2.5 years; SD = 7.4), and 48.4 months at the two-year follow-up (~4 years; SD = 7.5). More boys (n = 101; 55%) than girls participated, which may reflect the higher proportion of boys presenting with behaviour problems than girls from the age of two years (Sadler et al., 2018). Children's ethnicity, as reported by their primary caregivers, was 6% Asian, 8% Black, 20% Mixed heritage, 62% white, and 4% Other. This sample is more ethnically diverse than the overall UK population but likely reflects a blend of the higher levels of diversity within the four participating London Boroughs (percentage of two-year-olds recorded as having white ethnicity as follows: Barking and Dagenham: 39%; Camden: 51%; Hillingdon: 45%; Islington: 52%; Census, 2011a) and the lower levels of diversity in the three other sites (percentage of 0-4 year olds recorded as having white ethnicity as follows: Hertfordshire: 77%; Oxfordshire: 78%; Peterborough: 72%; Census, 2011a).

Children's primary caregivers were predominantly female (95%) and all were the child's biological or adoptive mother (95%) or father (5%). At baseline, 81% of participating caregivers were married or living with a partner. This is marginally higher than is seen in the general population of families with children aged 0-17 (78%; Office for National Statistics, 2017), possibly because relationship dissolution is lower in families with very young children. A higher proportion of caregivers held a graduate-level qualification (60%) than we would expect to see in the general population (41%; Census, 2011b). This disparity is not dissimilar to other studies as low parental education is a known barrier for the recruitment and retention of families with young children in clinical research (Robinson et al., 2016). Compared to 67% of women with a child aged 0-2 in employment in the general population (Office for National Statistics, 2019), only 55% of caregivers were currently employed or on paid parental leave, with the remainder looking after the home and family (41%) or in full-time education (4%).

Table 33

Sociodemographic Characteristics of the Full HSHS Sample and the Subsample Included in the Current Analysis

	Full HSHS sample			Current (sub)sample		
	Trial arm			Trial arm		
Characteristic	VIPP-SD (*n* = 151)	UC (*n* = 149)	All (*n* = 300)	VIPP-SD (*n* = 87)	UC (*n* = 98)	All (*n* = 185)
Child characteristic						
Sex (female), *n* (%)	75 (50)	62 (42)	137 (46)	40 (46)	44 (45)	84 (45)
Age at baseline (months), mean (SD)	22.8 (6.8)	23.2 (6.5)	23.0 (6.7)	22.9 (7.1)	23.5 (6.6)	23.2 (6.8)
Age at 5MFU (months), mean (SD)	29.6 (7.6)	28.8 (6.6)	29.2 (7.1)	29.6 (8.1)	28.9 (6.8)	29.1 (7.4)
Age at 2YFU (months), mean (SD)	47.9 (7.5)	47.8 (7.1)	47.8 (7.3)	48.5 (7.9)	48.2 (7.2)	48.4 (7.52)
Ethnicity, *n* (%)						
Asian	9 (6)	8 (5)	17 (6)	5 (6)	6 (6)	11 (6)
Black	3 (2)	15 (10)	18 (6)	3 (3)	12 (12)	15 (8)
Mixed ethnicity	36 (24)	25 (17)	61 (20)	21 (24)	16 (17)	37 (20)
White	100 (66)	94 (63)	194 (65)	55 (63)	59 (60)	114 (62)
Other	3 (2)	7 (5)	10 (3)	3 (4)	5 (5)	8 (4)
Primary caregiver characteristic						
Sex (female), *n* (%)	143 (95)	144 (97)	287 (96)	81 (93)	94 (96)	175 (95)

	Full HSHS sample			Current (sub)sample		
	Trial arm			Trial arm		
Characteristic	VIPP-SD (*n* = 151)	UC (*n* = 149)	All (*n* = 300)	VIPP-SD (*n* = 87)	UC (*n* = 98)	All (*n* = 185)
Age at baseline (years), mean (SD)	33.7 (5.6)	34.7 (5.9)	34.2 (5.8)	33.6 (5.5)	34.3 (6.1)	34.0 (5.8)
Parental status, *n* (%)						
Biological or adoptive mother	144 (95)	144 (97)	288 (96)	82 (94)	94 (96)	176 (95)
Biological or adoptive father	7 (5)	5 (3)	12 (4)	5 (6)	4 (4)	9 (5)
Ethnicity, *n* (%)						
Asian	15 (10)	16 (11)	31 (10)	8 (9)	11 (12)	19 (10)
Black	3 (2)	15 (10)	18 (6)	2 (3)	13 (13)	15 (8)
Mixed ethnicity	11 (7)	11 (7)	22 (8)	8 (9)	8 (8)	16 (9)
White	114 (75)	103 (69)	217 (72)	62 (71)	63 (64)	125 (68)
Other	8 (6)	4 (3)	12 (4)	7 (8)	3 (3)	10 (5)
Relationship status, *n* (%)						
Married/civil partnership/cohabiting	128 (85)	127 (85)	255 (85)	70 (81)	80 (82)	150 (81)
Single	12 (8)	17 (11)	29 (10)	8 (9)	13 (13)	21 (11)
In relationship but not cohabiting	10 (7)	1 (1)	11 (4)	8 (9)	1 (1)	9 (5)
Divorced/widowed/legally separated	1 (1)	4 (3)	5 (2)	1 (1)	4 (4)	5 (3)
Employment status, *n* (%)						

	Full HSHS sample			Current (sub)sample		
	Trial arm			Trial arm		
Characteristic	VIPP-SD (*n* = 151)	UC (*n* = 149)	All (*n* = 300)	VIPP-SD (*n* = 87)	UC (*n* = 98)	All (*n* = 185)
Employed	86 (57)	76 (51)	162 (54)	47 (54)	46 (47)	93 (50)
Looking after home and children	56 (37)	56 (37)	112 (37)	35 (40)	40 (41)	75 (41)
Paid parental leave	6 (4)	10 (7)	16 (5)	3 (3)	7 (7)	10 (5)
Full-time student	3 (2)	7 (5)	10 (4)	2 (2)	5 (5)	7 (4)
Highest qualification, *n* (%)						
Graduate	92 (61)	99 (66)	191 (64)	52 (60)	60 (61)	112 (60)
A-level/NVQ/BTEC	42 (28)	36 (24)	78 (26)	23 (26)	26 (27)	49 (27)
GCSE or lower	17 (11)	14 (10)	31 (10)	12 (14)	12 (12)	24 (13)

Note. VIPP-SD = Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline; UC = usual care.

5.4.3.3 Descriptive analysis of study variables

Descriptive statistics including means, standard deviations, and measures of skew and kurtosis were generated for key study variables for both the full HSHS sample and the subsample for the current study (see Table 34). Compared to the full sample, the current sample showed slightly higher levels of behaviour problems measured on the CBCL at all time points.

5.4.3.4 Correlational analysis of study variables

Correlational analyses were run for key study variables to explore the stability of children's behaviour over time (see Table 35). There was a large, positive association between children's behaviour problems at baseline and at 5-month follow-up ($r = .76$, $p < .001$), and a large, positive association between behaviour problems at 5-month follow-up and 24-month follow-up ($r = .57$, $p < .001$). This suggests that there is some stability over time of children's behaviour problems measured using the CBCL. There was a large, positive association between primary caregivers' sensitivity measured at baseline and 5-month follow-up ($r = .63$, $p < .001$). This suggests that there is some stability over time of parenting behaviours related to sensitivity measured using the Erickson scales.

Table 34

Descriptive Statistics of Study Variables for the Full HSHS Sample and Subsample Included in the Current Analyses

	Full sample					Current (sub)sample				
Variable	*N*	*M*	*SD*	Skewness	Kurtosis	*N*	*M*	*SD*	Skewness	Kurtosis
Baseline assessment										
CBCL baseline – total	300	41.68	21.39	0.94	0.85	178	42.36	20.86	0.85	0.50
Parental supportiveness	.	.	.	.	.	178	4.46	1.58	-0.19	-0.82
Parental intrusiveness	.	.	.	.	.	178	4.07	1.62	-0.15	-0.89
5-month assessment										
CBCL 5-month follow-up – total	285	34.90	20.90	0.89	0.49	177	36.04	20.87	0.85	0.41
Parental supportiveness	.	.	.	.	.	178	4.75	1.40	-0.44	-0.31
Parental intrusiveness	.	.	.	.	.	178	4.09	1.55	-0.09	-0.72
24-month assessment										
CBCL 24-month follow-up – total	285	32.97	23.63	1.59	3.40	174	33.43	23.49	1.45	2.33
SS Disruptive story themes	256	5.00	4.10	1.06	0.66	157	4.94	3.93	1.00	0.45
SS Prosocial story themes	256	3.68	1.93	-0.04	-0.74	157	3.65	1.84	-0.13	-0.66
SS Narrative coherence	256	10.38	2.78	-0.37	-0.24	157	10.38	2.76	-0.46	-0.22
SS Observed behavioural dysregulation	256	6.15	3.73	0.62	-0.23	157	6.01	3.75	0.67	-0.29

Note. CBCL = Child Behavior Checklist; Parental supportiveness and intrusiveness = Erickson scales; SS = Story stem measure

Table 35

Bivariate Correlations Between Measures of Parental Sensitivity, Child Behaviour, and Children's Narrative Representations

Variable	1	2	3	4	5	6	7	8	9
1. Parental sensitivity - baseline	-								
2. Parental sensitivity – 5-month	.63**	-							
3. CBCL Baseline	-.25**	-.18*	-						
4. CBCL 5-month	-.23**	-.15	.76**	-					
5. CBCL 2-year	-.07	-.02	.47**	.57**	-				
6. SS Disruptive story themes	-.08	.01	.22**	.10	.13	-			
7. SS Prosocial story themes	.10	.05	.07	.00	-.03	-.07	-		
8. SS Narrative coherence	.26**	.15	-.05	-.08	-.06	-.16*	.67**	-	
9. SS Observed behavioural dysregulation	-.26**	-.15	.11	.12	.17*	.40**	-.31**	-.28**	-

Note. * $p < .05$, ** $p < .01$

CBCL = Child Behavior Checklist; Parental sensitivity = Erickson scales; SS = Story stem measure

Main analysis

5.4.4 Aim 1: To explore associations between early parental sensitivity and children's later narratives

For the first aim, four separate regression models were conducted for each of the scales derived from the children's story stem data (narrative coherence, behavioural dysregulation, and expression of destructive and prosocial themes). Covariates included in these models were informed by the analysis of these variables in Chapters 3 and 4.

5.4.4.1 Checking the assumptions of multiple linear regression

Prior to conducting the regression analyses, preliminary tests were run to ensure the data met the assumptions of a hierarchical regression analysis. The findings of these checks and tests are summarised here for all regression models included in this analysis. Under the central limit theorem, we can assume with the size of this sample (N = 154 for cases with complete data across time points) that our betas are from a normally distributed sampling distribution (Field, 2013). A rough rule of thumb for adequate sample size is that we would need 10-15 cases of data per predictor included in the model (Peduzzi et al., 1996). For this sample, with up to five predictors (child age, child gender, child behaviour, trial arm, parental sensitivity), it is recommended the sample includes at least 75 participants, which the current sample size surpasses.

The data was firstly examined to explore associations between included variables in each regression using correlation matrices. No variables were highly associated (above 0.8 or 0.9; Field, 2013). The variance inflation factor (VIF) and tolerance statistic, which explore whether the included predictors had a strong linear relationship with the other predictors, were also used to assess multicollinearity. All variance inflation factors for all predictors were below 10 (Myers, 1990) and all tolerance statistics were above 0.2 (Menard, 1995), meaning there was no evidence of multicollinearity.

Data was checked for outliers, normality, non-linearity, heteroscedasticity and independence of residuals using diagnostic statistics and by inspecting normality probability plots (P-P plots) and scatterplots.

Firstly, plots of standardised residuals against standardised predicted values were examined. Dots appeared to be randomly and evenly dispersed throughout the plots, indicating that the assumptions of linearity and homoscedasticity were met. All standardised

residuals fell between -3.30 and 3.30 apart from one case in the disruptive themes model which was 3.54 (Tabachnick & Fidell, 2014). Further tests for outliers were conducted. Cook's distance measures the overall influence of a case on the model, with values greater than 1 being a cause for concern (Cook & Weisberg, 1982). No values were greater than 1 in these models (0.00-0.09). Leverage statistics (or hat values) gauge the influence of the observed value of the dependent variable over the independent variables (Field, 2013). These values can fall between 0 and 1. Using the number of predictors and sample size, cut-off points were calculated to identify any cases having undue influence. No cases exceeded three times the average leverage value (Stevens, 2002). Finally, Mahalanobis distances were used to detect multivariate outliers. This test measures the influence of a case by examining the distance of cases from the means of the predictor variables. All values fell below the critical values for Mahalanobis distances of each model.

Partial plots were also examined to detect outliers and heteroscedasticity. There were no obvious outliers on the plots, and the clouds of dots were evenly spaced around the line, indicating homoscedasticity. To test the normality of residuals, histograms and normal probability plots were inspected. Distributions appeared to be normal for each included regression, with the histograms being fairly symmetrical and bell-shaped. P-P plots were examined to assess any deviations from normality. For the regressions exploring prosocial themes, narrative coherence, and observed dysregulation, the dots lay close to the diagonal line indicating a normal distribution. For disruptive themes, the dots mostly stayed on the diagonal line, although there was a slight deviation from the line indicating it may be non-normal.

The assumption of independence of errors was assessed using the Durbin-Watson test, which tests whether adjacent residuals are correlated. Values for this statistic vary between 0 and 4, with a value of 2 meaning the residuals are uncorrelated (Field, 2013). The Durbin-Watson statistic for the data were: narrative coherence: 2.03, behavioural dysregulation: 2.14, disruptive story themes: 1.79, and prosocial story themes: 1.91.

This data generally met the assumptions for a multiple linear regression.

5.4.4.2 Exploring associations between early parental sensitivity and children's story stem representations two years later

Following the assumption checks, multiple regression analyses were conducted to explore whether early parental sensitivity (measured when children were between 12 and 36 months) predicted aspects of children's narrative representations two years later. Narrative coherence, behavioural dysregulation, disruptive story themes, and prosocial themes were explored separately. Effect size were calculated for each of the models using Cohen's f^2 (calculated by $R^2/(1-R^2)$) where, as a guide, 0.02 = small, 0.15 = medium, and 0.35 = large effects (Cohen, 1969, 1988).

5.4.4.3 Model 1: Does early parental sensitivity predict children's narrative coherence?

First, a multiple regression was conducted to explore whether baseline parental sensitivity, measured using observational coding of parent-child interactions when children were 12 to 36 months old, was associated with the coherence of children's storytelling two years later.

Total score of narrative coherence was the dependent variable. Child age and gender were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered second. Observed parental sensitivity, as measured using the Erickson scales of supportiveness and intrusiveness, was entered last. Table 36 displays the results of this analysis.

After controlling for the effects of potential confounding variables (child age, child gender, and trial arm), higher parental sensitivity at baseline was associated with greater narrative coherence in children's stories two years later when they were three and four years old (standardised $\beta = .21$, $p = .021$). The overall model fit was $R^2 = 0.21$ (accounting for 21% of the variance). This model represents a medium effect size, $f^2 = 0.28$.

Table 36

Linear Model of Predictors of Narrative Coherence

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	6.28 (3.27, 9.29)	1.53		4.12	.000
	Child age	0.13 (0.07, 0.18)	0.03	.33	4.41	.000
	Child gender	-1.31 (-2.12, -0.51)	0.41	-.24	-3.22	.002
2	(Constant)	5.76 (2.75, 8.78)	1.53		3.77	.000
	Child age	0.13 (0.07, 0.19)	0.03	.33	4.53	.000
	Child gender	-1.28 (-2.08, -0.49)	0.40	-.23	-3.18	.002
	Trial arm	0.87 (0.07, 1.67)	0.41	.16	2.16	.033
3	(Constant)	6.03 (3.05, 9.01)	1.51		4.00	.000
	Child age	0.12 (0.06, 0.17)	0.03	.30	4.12	.000
	Child gender	-1.12 (-1.92, -0.33)	0.40	-.21	-2.79	.006
	Trial arm	0.85 (0.07, 1.64)	0.40	.16	2.14	.034
	Parental sensitivity	0.50 (0.08, 0.92)	0.21	.18	2.34	.021

Note. R^2 = .16 for Step 1, ΔR^2 = .03 for Step 2 (*p*s = .033), ΔR^2 = .03 for Step 3 (*p*s = .021)

N = 157

5.4.4.4 Model 2: Does early parental sensitivity predict children's behavioural dysregulation?

Second, a multiple regression was conducted to explore whether early parental sensitivity, measured using observational coding of parent-child interactions when children were 12 to 36 months old, was associated with children's observed behavioural dysregulation during the completion of a story stem task at three to five years old.

Total score of behavioural dysregulation was the dependent variable. Child age and gender were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered next. Observed parental sensitivity, as measured using the Erickson scales of supportiveness and intrusiveness, was entered last. Table 37 displays the results of this analysis.

After controlling for the effects of potential confounding variables (child age, child gender, and trial arm), higher parental sensitivity at baseline was associated with lower behavioural dysregulation during the storytelling task two years later when children were three and four years old (standardised $\beta = -.22$, $p = .006$). The overall model fit was $R^2 = 0.13$ (accounting for 13% of the variance). This model represents a medium effect size, $f^2 = 0.15$.

Table 37

Linear Model of Predictors of Behavioural Dysregulation

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.71 (-1.58, 6.99)	2.17		1.25	.21
	Child age	-0.00 (-0.08, 0.08)	0.04	-.00	-0.02	.98
	Child gender	2.18 (1.03, 3.32)	0.58	.29	3.76	.000
2	(Constant)	2.80 (-1.55, 7.16)	2.20		1.27	.21
	Child age	-0.00 (-0.08, 0.08)	0.04	-.00	-0.03	.98
	Child gender	2.17 (1.02, 3.32)	0.58	.29	3.73	.000
	Trial arm	-0.16 (-1.31, 1.00)	0.59	-.02	-0.27	.79
3	(Constant)	2.35 (-1.92, 6.62)	2.16		1.09	.28
	Child age	0.02 (-0.06, 0.10)	0.04	.04	0.45	.65
	Child gender	1.90 (0.76, 3.04)	0.58	.26	3.29	.001
	Trial arm	-0.12 (-1.26, 1.01)	0.57	-.02	-0.22	.83
	Parental sensitivity	-0.86 (-1.46, -0.25)	0.31	-.22	-2.80	.006

Note. R^2 = .09 for Step 1, ΔR^2 = .000 for Step 2 (*p*s = .79), ΔR^2 = .05 for Step 3 (*p*s = .006)

N = 157

5.4.4.5 Model 3: Does early parental sensitivity predict children's expression of disruptive story themes?

Third, a multiple regression was conducted to explore whether early parental sensitivity, measured using observational coding of parent-child interactions when children were 12 to 36 months old, was associated with children's expression of disruptive story themes in the story stem task at three to five years old.

Total score of disruptive story themes was the dependent variable. Child age and gender were included in the first step of the model. Child behaviour was entered second. Intervention arm (VIPP-SD or usual care) was entered next. Observed parental sensitivity, as measured using the Erickson scales of supportiveness and intrusiveness, was entered last. Table 38 displays the results of this analysis.

After controlling for the effects of potential confounding variables (child age, child gender, child behaviour, and trial arm), parental sensitivity at baseline was not a statistically significant predictor of children's expression of disruptive story themes in their storytelling two years later when they were three and four years old (standardised $\beta = -.001$, $p = .99$). The overall model fit was $R^2 = 0.11$ (accounting for 11% of the variance). This model represents a small effect size, $f^2 = 0.12$.

Table 38

Linear Model of Predictors of Children's Expression of Disruptive Story Themes

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	-1.16 (-5.66, 3.35)	2.28		-0.51	.61
	Child age	0.06 (-0.02, 0.15)	0.04	.12	1.49	.14
	Child gender	1.96 (0.76, 3.16)	0.61	.25	3.23	.002
2	(Constant)	-1.74 (-6.19, 2.71)	2.25		-0.77	.44
	Child age	0.04 (-0.05, 0.13)	0.04	.07	0.90	.37
	Child gender	2.09 (0.91, 3.28)	0.60	.27	3.49	.001
	Child behaviour	0.04 (0.01, 0.07)	0.02	.20	2.52	.013
3	(Constant)	-1.89 (-6.42, 2.64)	2.29		-0.83	.41
	Child age	0.04 (-0.05, 0.13)	0.04	.07	0.90	.37
	Child gender	2.10 (0.91, 3.29)	0.60	.27	3.49	.001
	Child behaviour	0.04 (0.01, 0.07)	0.02	.20	2.54	.012
	Trial arm	0.23 (-0.97, 1.44)	0.61	.03	0.38	.70
4	(Constant)	-1.90 (-6.44, 2.65)	2.30		-0.82	.41
	Child age	0.04 (-0.05, 0.13)	0.05	.07	0.87	.39
	Child gender	2.10 (0.88, 3.31)	0.62	.27	3.40	.001
	Child behaviour	0.04 (0.01, 0.07)	0.02	.20	2.42	.017
	Trial arm	0.23 (-0.97, 1.44)	0.61	.03	0.38	.70
	Parental sensitivity	-0.00 (-0.67, 0.67)	0.34	-.00	-0.01	.99

Note. R^2 = .08 for Step 1, ΔR^2 = .04 for Step 2 (*p*s = .013, ΔR^2 = .001 for Step 3 (*p*s = .70), ΔR^2 = .00 for Step 4 (*p*s = .99)
N = 157

5.4.4.6 Model 4: Does early parental sensitivity predict children's expression of prosocial story themes?

Fourth, a multiple regression was conducted to explore whether early parental sensitivity, measured using observational coding of parent-child interactions when children were 12 to 36 months old, was associated with children's expression of prosocial story themes in the story stem task two years later.

Total score of prosocial story themes was the dependent variable. Child age and gender were included in the first step of the model. Child behaviour was entered second. Intervention arm (VIPP-SD or usual care) was entered next. Observed parental sensitivity, as measured using the Erickson scales of supportiveness and intrusiveness, was entered last. Table 39 displays the results of this analysis.

After controlling for the effects of potential confounding variables (child age, child gender, child behaviour, and trial arm), parental sensitivity at baseline was not a statistically significant predictor of children's expression of prosocial story themes in their storytelling two years later when they were three and four years old (standardised $\beta = -.01$, $p = .85$). While the overall model fit was $R^2 = 0.26$ (accounting for 26% of the variance) and represents a medium effect size ($f^2 = 0.19$), this effect was largely driven by the model's covariates, rather than parental sensitivity.

Table 39

Linear Model of Predictors of Children's Expression of Prosocial Story Themes

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	0.93 (-1.00, 2.85)	0.97		0.95	.34
	Child age	0.09 (0.06, 0.13)	0.02	.36	5.14	.000
	Child gender	-1.19 (-1.71, -0.68)	0.26	-.32	-4.59	.000
2	(Constant)	0.99 (-0.95, 2.93)	0.98		1.01	.31
	Child age	0.10 (0.06, 0.13)	0.02	.37	5.14	.000
	Child gender	-1.21 (-1.72, -0.69)	0.26	-.33	-4.62	.000
	Child behaviour	-0.00 (-0.02, 0.01)	0.01	-.05	-0.65	.51
3	(Constant)	0.59 (-1.35, 2.52)	0.98		0.60	.55
	Child age	0.10 (0.06, 0.13)	0.02	.37	5.21	.000
	Child gender	-1.19 (-1.70, -0.68)	0.26	-.32	-4.63	.000
	Child behaviour	-0.00 (-0.02, 0.01)	0.01	-.02	-0.28	.78
	Trial arm	0.63 (0.12, 1.15)	0.26	.17	2.43	.016
4	(Constant)	0.58 (-1.37, 2.52)	0.98		0.59	.56
	Child age	0.10 (0.06, 0.14)	0.02	.38	5.08	.000
	Child gender	-1.20 (-1.72, -0.68)	0.26	-.33	-4.55	.000
	Child behaviour	-0.00 (-0.02, 0.01)	0.01	-.02	-0.32	.75
	Trial arm	0.63 (0.11, 1.15)	0.26	.17	2.41	.017
	Parental sensitivity	-0.03 (-0.31, 0.26)	0.15	-.01	-0.19	.85

Note. R^2 = .23 for Step 1, ΔR^2 = .002 for Step 2 (*p*s = .51, ΔR^2 = .030 for Step 3 (*p*s = .016), ΔR^2 = .00 for Step 4 (*p*s = .85)
N = 157

To summarise the findings of Aim 1, baseline parental sensitivity (at 12 to 36 months old) predicted children's narrative coherence and behavioural dysregulation (at three to five years old) but not their use of disruptive or prosocial themes. Now, I explore whether effects of the VIPP-SD intervention on children's narrative coherence was mediated by post-treatment parental sensitivity.

5.4.5 Aim 2: To explore whether intervention effects on narrative coherence were mediated by parental sensitivity.

In the previous chapter, results suggested that children's narrative coherence differed based on their group allocation, in that children whose parents received the VIPP-SD intervention exhibited higher levels of narrative coherence than those in the control group. The current study provides access to a pre- and post-treatment observational assessment of parenting for a subsample of participating families (n = 154, 60% of the previous chapter's sample; 51% of the full HSHS sample). This observational measure offers one potential mediating mechanism of the relationship seen in the previous chapter. To explore this possibility further, I undertook analysis in three main steps.

1. **Regressions:** First, I constructed four regression models:
    a. The first explored the association between group allocation and children's narrative coherence at the 24-month follow-up, as measured using the story stems.
    b. The second explored associations between group allocation and parents' 5-month sensitivity score, as measured using the Erickson scales.
    c. The third and fourth models formed further exploratory analysis whereby associations between group allocation and the two Erickson subscales (supportiveness and intrusiveness) were examined separately.
2. **Independent samples *t*-tests:** Second, I explored mean group differences on outcome (narrative coherence; story stems) and mediator (parental sensitivity; Erickson scales) variables using *t*-tests.
3. **Mediation:** Third, I ran a simple mediation analysis to explore whether parental sensitivity measured post-intervention mediated any relationship between group allocation and children's narrative coherence at the 24-month follow-up.

5.4.5.1 Regression models: Exploring associations between group allocation and child (narrative coherence) and parent (parental sensitivity) outcomes

5.4.5.1.1 Regression 1: Group allocation as a predictor of children's narrative coherence

First, a multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with children's narrative coherence, measured through a story stem task completed when children were aged three to four years old. Total

narrative coherence score was the dependent variable. Child age and gender were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. After controlling for the effects of potential confounding variables (child age and child gender), VIPP-SD group allocation was associated with higher narrative coherence at the 24-month follow-up (standardised β = .16, p = .033). The overall model fit was R^2 = 0.19 (accounting for 19% of the variance) and represents a medium effect size (f^2 = 0.23). The magnitude of effect of the programme on narrative coherence was similar to that found in the full sample of children with story stem data, reported in the previous chapter (N = 256; standardised β = .19). Note, this data is not tabulated here as it has already been presented as part of the analysis in Aim 1 of this section (Table 36).

5.4.5.1.2 Regression 2: Group allocation as a predictor of parents' 5-month sensitivity scores

Second, a multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' overall sensitivity scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson scales. Mean sensitivity score was the dependent variable. Child age, child gender, and baseline sensitivity scores were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. Table 40 displays the results of this analysis. After controlling for the effects of potential confounding variables, group allocation was not associated with parental sensitivity at the 5-month follow-up (standardised β = .05, p = .44), accounting for .003 (0.3%) of the model's variance. The overall model fit was R^2 = .38 (accounting for 38% of the variance). This model represents a large effect size, f^2 = 0.62, which is largely driven by the contribution of baseline sensitivity.

Table 40

Linear Model of Predictors of Five-Month (Post-Intervention) Parental Sensitivity

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	-0.12 (-0.65, 0.42)	0.27		-0.43	.67
	Child age	0.01 (-0.01, 0.03)	0.01	.08	1.24	.22
	Child gender	-0.07 (-0.30, 0.17)	0.12	-.04	-0.55	.58
	Baseline sensitivity	0.57 (0.44, 0.69)	0.06	.59	8.97	.000
2	(Constant)	-0.18 (-0.73, 0.38)	0.28		-0.62	.54
	Child age	0.01 (-0.01, 0.03)	0.01	.09	1.30	.20
	Child gender	-0.06 (-0.30, 0.17)	0.12	-.04	-0.53	.60
	Baseline sensitivity	0.57 (0.44, 0.69)	0.06	.59	8.93	.000
	Group allocation	0.09 (-0.14, 0.33)	0.12	.05	0.78	.44

Note. R^2 = .38 for Step 1, ΔR^2 = .003 for Step 2 (ps = .44)

N = 154

5.4.5.1.3 Exploratory analysis – Regression 3: Group allocation as a predictor of parents' 5-month supportiveness (Erickson subscale)

In further exploratory analysis, the supportiveness subscale of the Erickson scales was examined. A multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' supportiveness scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson subscale. Mean supportiveness score was the dependent variable. Child age, child gender, and baseline supportiveness score were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. Table 41 displays the results of this analysis. After controlling for the effects of potential confounding variables, group allocation was not associated with parental supportiveness at the 5-month follow-up (standardised β = .04, p = .56), accounting for .002 (0.2%) of the model's variance. The overall model fit was R^2 = .29 (accounting for 29% of the variance). This model represents a large effect size, $\hat{f}^2$ = 0.42, which is largely driven by the contribution of baseline supportiveness.

Table 41

Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Supportiveness (Erickson Subscale)

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.82 (1.83, 3.80)	0.50		5.65	.000
	Child age	0.01 (-0.02, 0.04)	0.01	.05	0.76	.45
	Child gender	-0.14 (-0.50, 0.22)	0.18	-.05	-0.76	.45
	Baseline supportiveness	0.44 (0.33, 0.56)	0.06	.52	7.37	.000
2	(Constant)	2.76 (1.75, 3.77)	0.51		5.41	.000
	Child age	0.01 (-0.02, 0.04)	0.01	.06	0.81	.42
	Child gender	-0.14 (-0.50, 0.23)	0.19	-.05	-0.75	.46
	Baseline supportiveness	0.44 (0.32, 0.56)	0.06	.52	7.30	.000
	Group allocation	0.11 (-0.26, 0.47)	0.18	.04	0.59	.56

Note. R^2 = .29 for Step 1, ΔR^2 = .002 for Step 2 (ps = .56)

N = 154

5.4.5.1.4 Exploratory analysis – Regression 4: Group allocation as a predictor of parents' 5-month intrusiveness (Erickson subscale)

In further exploratory analysis, the intrusiveness subscale of the Erickson scales was examined. A multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' intrusiveness scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson subscale. Mean intrusiveness score was the dependent variable. Child age, child gender, and baseline intrusiveness score were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. Table 42 displays the results of this analysis. After controlling for the effects of potential confounding variables, group allocation was not associated with parental intrusiveness at the 5-month follow-up (standardised β = -.06, p = .34), accounting for .004 (0.4%) of the model's variance. The overall model fit was R^2 = .41 (accounting for 41% of the variance). This model represents a large effect size, f^2 = 0.70, which is largely driven by the contribution of baseline intrusiveness.

Table 42

Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Intrusiveness (Erickson Subscale)

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.13 (1.13, 3.12)	0.50		4.23	.000
	Child age	-0.02 (-0.05, 0.00)	0.01	-.11	-1.69	.09
	Child gender	0.07 (-0.31, 0.45)	0.19	.02	0.36	.72
	Baseline intrusiveness	0.58 (0.46, 0.70)	0.06	.61	9.48	.000
2	(Constant)	2.24 (1.22, 3.26)	0.52		4.33	.000
	Child age	-0.03 (-0.05, 0.00)	0.01	-.11	-1.75	.08
	Child gender	0.06 (-0.32, 0.45)	0.19	.02	0.33	.74
	Baseline intrusiveness	0.58 (0.46, 0.70)	0.06	.61	9.48	.000
	Group allocation	-0.18 (-0.57, 0.20)	0.19	-.06	-0.95	.34

Note. R^2 = .41 for Step 1, ΔR^2 = .004 for Step 2 (ps = .34)

N = 154

5.4.5.2 Independent samples t-tests: Exploring group differences on mediator (parental sensitivity) and outcome (narrative coherence) variables

Independent samples t-tests were used to explore differences in outcomes by group allocation for those families with complete data across the two years of the study. There was a difference in the sample sizes of the usual care (n = 86) and intervention (n = 68) groups. The distributional assumptions necessary for an independent t-test have already been discussed in section 5.4.4.1. Levene's test, a test of the homogeneity of variance, was also run. Standardised effect sizes (Cohen's d) were computed using the difference in means between the two groups and dividing this by the standard deviation of the control arm. Cohen's d effect sizes can be defined as 0.20 = small, 0.50 = medium, and 0.80 = large (Cohen, 1988).

5.4.5.2.1 Group differences on parental sensitivity (mediator, *M*)

For parental sensitivity, Levene's test was not significant (F = 0.46, p = .50) indicating equal variance. At the five-month follow-up assessment, caregivers in the VIPP-SD group (M = 0.16, SE = 0.10) scored more highly for parental sensitivity than caregivers in the usual care group (M = 0.05, SE = 0.10). This difference, -0.10, 95% CI [-0.40, 0.19], was not statistically significant, t(152) = -0.70, p = .49, and represented a small-sized effect, Cohen's d = 0.11.

For information, Table 43 presents the ns, means, and standard deviations of baseline and 5-month supportiveness and intrusiveness scale scores, split by group allocation.

Table 43

Baseline and Five-Month Parental Supportiveness and Intrusiveness by Trial Arm

	Trial arm					
	VIPP SD			Usual care		
	N	Supportiveness M (SD)	Intrusiveness M (SD)	N	Supportiveness M (SD)	Intrusiveness M (SD)
Baseline	71	4.69 (1.55)	3.99 (1.67)	86	4.53 (1.54)	3.89 (1.57)
5-month	68	4.99 (1.23)	3.86 (1.45)	86	4.81 (1.40)	3.99 (1.59)

N = 157 at baseline and N = 154 at 5-month follow-up

5.4.5.2.2 Group differences on narrative coherence (outcome, *Y*)

For narrative coherence, Levene's test was not significant (F = 0.04, p = .85) indicating equal variance. At the 24-month follow-up assessment, children from the VIPP-SD

group (*M* = 10.90, *SE* = 0.29) scored more highly for narrative coherence than children in the usual care group (*M* = 10.03, *SE* = 0.34). This difference, -0.86, 95% CI [-1.73, 0.01], t(152) = -1.96, *p* = .05, represented a small-sized effect, Cohen's *d* = 0.32. This is comparable to the effect size reported on the full sample in the previous chapter (*N* = 256; *d* = .39).

5.4.5.3 Exploring mediation effects of parental sensitivity on the association between group allocation and children's narrative coherence

A mediation model was run using the *PROCESS* v3.5 (Hayes, 2018) extension in SPSS 25. This analysis calculated the effects for five paths:

1. *a* path: The simple regression of parental sensitivity predicted by group allocation.
2. *b* path: The regression of narrative coherence predicted by parental sensitivity.
3. *c path:* The regression of narrative coherence predicted by group allocation without parental sensitivity in the model.
4. *c'* path: The regression of narrative coherence predicted by group allocation with parental sensitivity in the model.
5. *ab* path: The indirect effect of group allocation on narrative coherence.

In this analysis, data from 154 participating families were included. Baseline parental sensitivity, and child age and gender were entered as covariates. The relationship between group allocation and children's narrative coherence was not mediated by parental sensitivity. The results are presented in Figure 16, with the effect for each path displayed alongside the standard error in brackets. The bootstrapped unstandardised indirect effect of group allocation on narrative coherence through parental sensitivity was not statistically significant, *b* = -0.02 (0.05), BCa CI [-0.160, 0.066], with little evidence of an indirect effect via sensitivity.

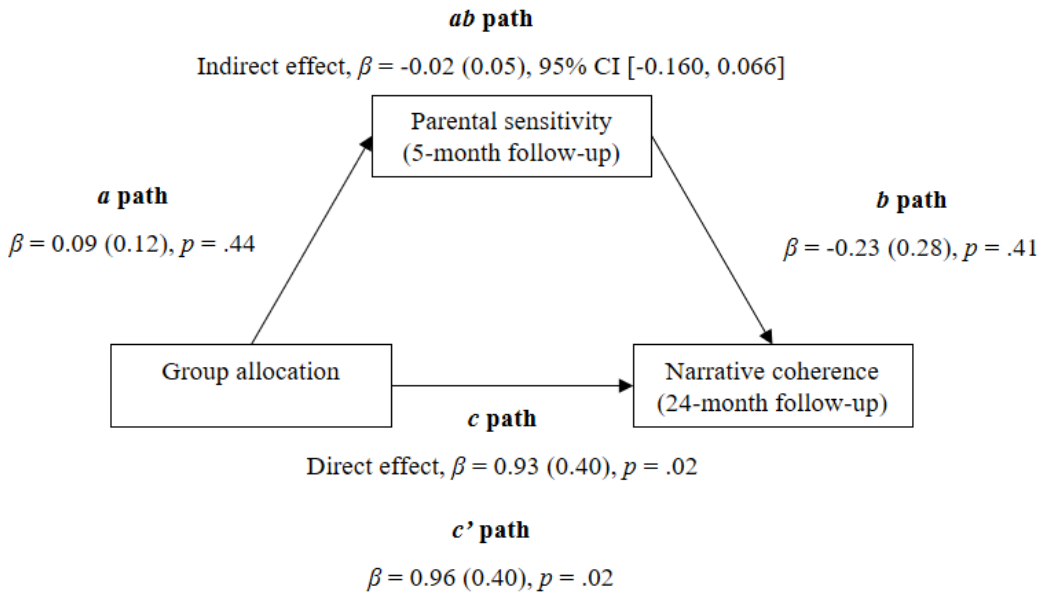


Figure 16. Simple mediation model of intervention effects on children's narrative coherence mediated by parental sensitivity at 5-month follow-up.

5.4.5.4 Aim 2: Summary of findings

Before moving onto Aim 3, to summarise the findings of Aim 2, group allocation was associated with higher narrative coherence on the story stems at the 24-month follow-up but not parental sensitivity at the 5-month follow-up. Mean group differences were small for narrative coherence (d = .39) and very small for parental sensitivity (d = .11), both favouring the VIPP-SD group. There was no evidence of the mediating effect of parental sensitivity on the relationship between group allocation and children's narrative coherence.

5.4.6 Aim 3: To explore whether intervention effects on children's behaviour (as measured through the CBCL) were mediated by parental sensitivity.

Finally, the influence of parental sensitivity on the relationship between families' treatment allocation and children's behaviour problems was examined. A slightly larger sample was available for this analysis (N = 178), compared to that available in Aim 2 (N = 154), due to the fact this aim was not reliant on children's completion of the story stem battery at the 24-month follow-up, but rather parents' completion of the CBCL at the five-month follow-up. Because of this, some steps of analysis were repeated to ensure findings were reflective of this slightly larger sample. Analysis was undertaken in three main steps:

1. **Regressions:** First, I constructed four regression models:
    a. The first explored the association between group allocation and children's behaviour problems measured at the 5-month follow-up using the CBCL.
    b. The second explored associations between group allocation and parents' 5-month sensitivity score, as measured using the Erickson scales.
    c. The third and fourth models formed further exploratory analysis whereby associations between group allocation and the two Erickson subscales (supportiveness and intrusiveness) were examined separately.
2. **Independent samples *t*-tests:** Second, I explored mean group differences on outcome (child behaviour problems; CBCL) and mediator (parental sensitivity; Erickson scales) variables using *t*-tests.
3. **Mediation:** Third, I ran a mediation analysis to explore whether parental sensitivity measured post-intervention mediated any relationship between group allocation and children's behaviour problems at the 5-month follow-up.

5.4.6.1 Regression models: Exploring associations between group allocation and child (child behaviour, CBCL) and parent (parental sensitivity) outcomes

5.4.6.1.1 Regression 1: Group allocation as a predictor of children's behaviour (CBCL) at the 5-month follow-up (post-intervention)

First, a multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with children's post-treatment behaviour problems, measured using the parent-reported CBCL. Total score on the CBCL was the dependent variable. Child age and gender were included in the first step of the model. Child behaviour measured through the CBCL at baseline was entered second. Intervention arm (VIPP-SD or

usual care) was entered last. Table 44 displays the results of this analysis. After controlling for the effects of potential confounding variables (child age, child gender, and baseline child behaviour), group allocation was not associated with children's behaviour post-intervention (standardised β = -.07, p = .17) to a statistically significant degree. While the overall model fit was R^2 = 0.58 (accounting for 58% of the variance) and represents a large effect size (f^2 = 1.35), this effect was largely driven by baseline child behaviour, rather than the intervention.

For information, a post-hoc analysis was run to see whether similar associations between group allocation and child behaviour at post-intervention were present in the full HSHS sample (N = 285). Comparably sized effects were found in the full sample (standardised β = -.05, p = .14; see Appendix O for full output).

Table 44

Linear Model of Predictors of Post-Intervention (Five-Month Follow-Up) Child Behaviour (CBCL)

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	23.20 (8.94, 37.47)	7.23		3.21	.002
	Child age	0.59 (0.15, 1.04)	0.23	.20	2.63	.009
	Child gender	-0.58 (-6.71, 5.55)	3.11	-.01	-0.19	.85
2	(Constant)	1.56 (-8.33, 11.44)	5.01		0.31	.76
	Child age	0.08 (-0.22, 0.38)	0.15	.03	0.53	.60
	Child gender	0.39 (-3.68, 4.46)	2.06	.01	0.19	.85
	Baseline behaviour	0.75 (0.66, 0.85)	0.05	.76	14.92	.000
3	(Constant)	3.55 (-6.71, 13.81)	5.20		0.68	.50
	Child age	0.08 (-0.22, 0.38)	0.15	.03	0.51	.61
	Child gender	0.25 (-3.81, 4.31)	2.06	.01	0.12	.90
	Baseline behaviour	0.75 (0.65, 0.85)	0.05	.75	14.65	.000
	Group allocation	-2.88 (-6.97, 1.22)	2.08	-.07	-1.39	.17

Note. R^2 = .04 for Step 1, ΔR^2 = .54 for Step 2 (ps = .000), ΔR^2 = .01 for Step 3 (ps = .17)

N = 177

5.4.6.1.2 Regression 2: Group allocation as a predictor of parents' 5-month sensitivity scores

Second, a multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' overall sensitivity scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson scales. Mean sensitivity score was the dependent variable. Child age, child gender, and baseline parental sensitivity were entered in the third step. Intervention arm (VIPP-SD or usual care) was entered last. Table 45 displays the results of this analysis. After controlling for the effects of potential confounding variables (child age, child gender, and baseline sensitivity), group allocation was not associated with changes to caregivers' sensitivity (standardised β = .10, p = .098) to a statistically significant degree, accounting for .01 (1%) of the model's variance. While the overall model fit was R^2 = 0.42 (accounting for 42% of the variance) and represents a large effect size (f^2 = 0.71), this effect was largely driven by baseline parental sensitivity.

Table 45

Linear Model of Predictors of Five-Month (Post-Intervention) Parental Sensitivity

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	-0.22 (-0.73, 0.30)	0.26		-0.83	.41
	Child age	0.01 (-0.00, 0.03)	0.01	.10	1.60	.11
	Child gender	-0.07 (-0.29, 0.16)	0.11	-.03	-0.57	.57
	Baseline sensitivity	0.60 (0.48, 0.72)	0.06	.61	10.10	.000
2	(Constant)	-0.32 (-0.84, 0.20)	0.27		-1.21	.23
	Child age	0.01 (-0.00, 0.03)	0.01	.10	1.68	.10
	Child gender	-0.06 (-0.28, 0.16)	0.11	-.03	-0.52	.60
	Baseline sensitivity	0.60 (0.48, 0.71)	0.06	.61	10.10	.000
	Group allocation	0.19 (-0.03, 0.41)	0.11	.10	1.67	.098

Note. R^2 = .41 for Step 1, ΔR^2 = .01 for Step 2 (ps = .098)

N = 178

5.4.6.1.3 Exploratory analysis – Regression 3: Group allocation as a predictor of parents' 5-month supportiveness (Erickson subscale)

As part of further exploratory analysis, the supportiveness subscale of the Erickson scales was examined for this slightly larger sample (N = 178). A multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' supportiveness scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson subscale (see Table 46). Mean supportiveness score was the dependent variable. Child age, child gender, and baseline supportiveness score were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. After controlling for the effects of potential confounding variables, group allocation was not associated with parental supportiveness at the 5-month follow-up (standardised β = .09, p = .13) to a statistically significant degree, accounting for .009 (0.9%) of the model's variance. The overall model fit was R^2 = .34 (accounting for 34% of the variance). This model represents a large effect size, f^2 = 0.49, which is largely driven by the contribution of baseline supportiveness.

Table 46

Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Supportiveness (Erickson Subscale)

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.37 (1.44, 3.30)	0.47		5.01	.000
	Child age	0.02 (-0.01, 0.04)	0.01	.09	1.40	.16
	Child gender	-0.14 (-0.49, 0.21)	0.18	-.05	-0.79	.43
	Baseline supportiveness	0.49 (0.38, 0.60)	0.06	.55	8.66	.000
3	(Constant)	2.24 (1.30, 3.19)	0.48		4.69	.000
	Child age	0.02 (-0.01, 0.04)	0.01	.09	1.48	.14
	Child gender	-0.13 (-0.48, 0.21)	0.18	-.05	-0.75	.45
	Baseline supportiveness	0.48 (0.37, 0.60)	0.06	.54	8.58	.000
	Group allocation	0.26 (-0.08, 0.61)	0.17	.09	1.51	.13

Note. R^2 = .34 for Step 1, ΔR^2 = .01 for Step 2 (ps = .13)

N = 178

5.4.6.1.4 Exploratory analysis – Regression 4: Group allocation as a predictor of parents' 5-month intrusiveness (Erickson subscale)

As part of further exploratory analysis, the intrusiveness subscale of the Erickson scales was examined for this slightly larger sample (N = 178). A multiple regression was conducted to explore whether treatment allocation (VIPP-SD or usual care) was associated with parents' intrusiveness scores at the 5-month follow-up (i.e., post-intervention), measured using the Erickson subscale (see Table 47). Mean intrusiveness score was the dependent variable. Child age, child gender, and baseline intrusiveness score were included in the first step of the model. Intervention arm (VIPP-SD or usual care) was entered last. After controlling for the effects of potential confounding variables, group allocation was not associated with parental intrusiveness at the 5-month follow-up (standardised β = -.10, p = .10) to a statistically significant degree, accounting for .01 (1%) of the model's variance. The overall model fit was R^2 = .42 (accounting for 42% of the variance). This model represents a large effect size, f^2 = 0.71, which is largely driven by the contribution of baseline intrusiveness.

Table 47

Exploratory Analysis: Linear Model of Predictors of Five-Month (Post-Intervention) Parental Intrusiveness (Erickson Subscale)

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.13 (1.18, 3.07)	0.48		4.43	.000
	Child age	-0.02 (-0.05, 0.00)	0.01	-.10	-1.72	.09
	Child gender	0.07 (-0.29, 0.44)	0.18	.02	0.40	.69
	Baseline intrusiveness	0.58 (0.47, 0.70)	0.06	.61	10.14	.000
2	(Constant)	2.29 (1.33, 3.26)	0.49		4.70	.000
	Child age	-0.02 (-0.05, 0.00)	0.01	-.11	-1.78	.08
	Child gender	0.06 (-0.30, 0.42)	0.18	.02	0.34	.74
	Baseline intrusiveness	0.59 (0.47, 0.70)	0.06	.61	10.21	.000
	Group allocation	-0.30 (-0.65, 0.06)	0.18	-.10	-1.64	.10

Note. R^2 = .41 for Step 1, ΔR^2 = .01 for Step 2 (ps = .10)

N = 178

5.4.6.2 Exploring group differences on mediator (parental sensitivity) and outcome (child behaviour; CBCL) variables

Independent samples t-tests were used to explore differences in outcomes by group allocation, as a precursor to the mediation analysis. There was a difference in the sample sizes of the usual care (n = 98) and intervention (n = 81) groups. As this difference was relatively small no further action was taken. The distributional assumptions necessary for an independent t-test have already been discussed in section 5.4.4.1. Levene's test, a test of the homogeneity of variance, was also run. Standardised effect sizes (Cohen's d) were computed using the difference in means between the two groups and dividing this by the standard deviation of the control group. As a guide, thresholds for Cohen's d effect sizes are as follows: 0.20 = small, 0.50 = medium, and 0.80 = large (Cohen, 1988).

5.4.6.2.1 Group differences in parental sensitivity (mediator, *M*) and subscales at 5-month follow-up

For parental sensitivity, Levene's test was not significant (F = 2.12, p = .15) indicating equal variance. At the five-month follow-up assessment, caregivers in the VIPP-SD group (M = 0.12, SE = 0.10) scored more highly for parental sensitivity than caregivers in the usual care group (M = -0.09, SE = 0.10). This difference, -0.21, 95% CI [-0.50, 0.07], t(176) = -1.48, p = .14, represented a small-sized effect, Cohen's d = 0.23.

Table 48 presents descriptive statistics of baseline and 5-month supportiveness and intrusiveness scale scores, split by group allocation. Independent samples t-tests were conducted to examine each subscale separately. This revealed that caregivers in the VIPP-SD group (M = 4.94, SE = 0.14) scored more highly for parental supportiveness than caregivers in the usual care group (M = 4.59, SE = 0.15). This difference, -0.35, 95% CI [-0.77, 0.06], t(176) = -1.67, p = .097, represented a small-sized effect, Cohen's d = 0.25. Caregivers in the VIPP-SD group (M = 3.94, SE = 0.16) scored lower for parental intrusiveness than caregivers in the usual care group (M = 4.21, SE = 0.17). This difference, 0.27, 95% CI [-0.19, 0.73], t(176) = 1.17, p = .243, represented a small-sized effect, Cohen's d = 0.18.

Table 48

Baseline and Five-Month Parental Supportiveness and Intrusiveness by Trial Arm

	Trial arm					
	VIPP-SD			Usual care		
	N	Supportiveness *M* (*SD*)	Intrusiveness *M* (*SD*)	*N*	Supportiveness *M* (*SD*)	Intrusiveness *M* (*SD*)
Baseline	81	4.56 (1.58)	4.08 (1.67)	97	4.37 (1.58)	4.06 (1.58)
5-month	81	4.94 (1.27)	3.94 (1.43)	97	4.59 (1.49)	4.21 (1.64)

Note. VIPP-SD = Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline

N = 178

5.4.6.2.2 Group differences in child behaviour (CBCL) at 5-month follow-up (outcome, *Y*)

For child behaviour at post-intervention follow-up, Levene's test was not significant ($F = 1.28$, $p = .26$) indicating equal variance. At the five-month follow-up assessment, children in the VIPP-SD group ($M = 32.14$, $SE = 2.17$) scored lower for behaviour problems than children in the usual care group ($M = 39.34$, $SE = 2.19$). This difference, 7.21, 95% CI [1.07, 13.34], t(175) = 2.32, $p = .02$, represented a small-sized effect, Cohen's $d = 0.35$.

For completeness, the same analysis was conducted using the full HSHS sample ($N = 285$) to see if comparable differences in child behaviour post-intervention were found in the complete sample. A smaller mean difference (4.78, 95% CI [-0.08, 9.62], t(283) = 1.94, $p = .05$) and effect size ($d = 0.23$) were found in the full sample.

5.4.6.3 Exploring mediation effects of parental sensitivity on children's behaviour problems (as measured through the CBCL)

A mediation model was run using the *PROCESS* v3.5 (Hayes, 2018) extension in SPSS 25. In this analysis, data from 177 participating families were included. Child age, child gender, baseline child behaviour and baseline parental sensitivity were entered as covariates.

This analysis calculated the effects for five paths:

1. *a* path: The simple regression of parental sensitivity predicted by group allocation.
2. *b* path: The regression of child behaviour predicted by parental sensitivity.

3. *c path:* The regression of child behaviour predicted by group allocation without parental sensitivity in the model.
4. *c'* path: The regression of child behaviour predicted by group allocation with parental sensitivity in the model.
5. *ab* path: The indirect effect of group allocation on child behaviour.

The relationship between group allocation and children's post-intervention behaviour was not mediated by parental sensitivity. The results are presented in Figure 17, with the effect for each path displayed alongside the standard error in brackets. The bootstrapped unstandardised indirect effect of group allocation on children's behaviour at post-intervention through parental sensitivity was not statistically significant, $\beta = 0.09$ (0.27), BCa CI [-0.50, 0.68].

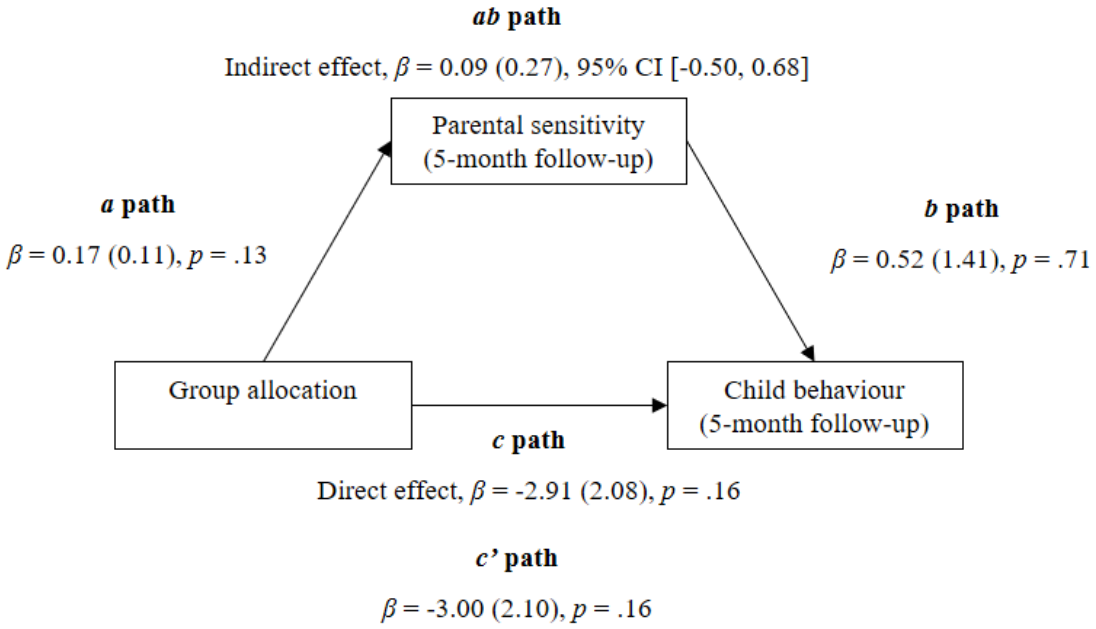


Figure 17. Mediation model of intervention effects on children's post-intervention behaviour mediated by parental sensitivity at 5-month follow-up.

5.4.6.4 Aim 3: Summary of findings

To summarise the findings of Aim 3, group allocation was not associated with child behaviour ($p = .17$) or parental sensitivity ($p = .098$) to a statically significant degree. Mean group differences were small for parental sensitivity ($d = 0.23$) and child behaviour ($d = 0.35$), both favouring the VIPP-SD group. There was no evidence of the mediating effect of parental sensitivity on the relationship between group allocation and children's behaviour problems.

5.5 Discussion

This study explored how children's early experiences of parental sensitivity may be associated with aspects of their later development. It considered this by first examining associations between early parental sensitivity and children's narrative representations, which aimed to access constructs related to their internal representations. Then, improvements in parental sensitivity post-intervention were measured to assess whether this acted as a mediating mechanism in the effects of a video-feedback parenting programme on children's narrative coherence and behaviour problems. The findings of this study are mixed. Early parental sensitivity, measured when children were one and two years old, was found to be associated with children's narrative coherence and their behavioural dysregulation during a story stem completion task at three and four years old but was not associated with children's expression of disruptive or prosocial story themes. Additionally, while there was evidence of the positive effects of the VIPP-SD programme on children's behaviour problems and their narrative coherence, effects of the programme on parental sensitivity were small and there was little to no evidence of parental sensitivity's mediating role in the programme effects on child outcomes. To my knowledge, this is the first study to include a measure of children's representations when examining the mechanisms of the VIPP-SD programme. Research findings are now discussed in detail before the strengths, limitations, and possible implications of this study are considered.

5.5.1 Overview and discussion of findings

In this section, I aim to better understand the current set of findings by contextualising them within the existing body of literature. Each aim of the study is considered in turn.

5.5.1.1 Aim 1: Exploring associations between early parental sensitivity and children's later narrative representations (as measured through the story stems)

In this study, the influences of parenting on children's internal representations were studied by using a narrative-based doll's house play story stem assessment battery. The findings are derived from four story components that broadly relate to the 'what' and the 'how' of children's narratives, as originally described by Sher-Censor and Oppenheim (2004). Children's expression of disruptive and prosocial story themes could be regarded as the 'what' of the narrative response, while the extent to which children exhibit coherence and regulation in their stories could be regarded as the 'how' of the narrative's construction. Each construct's relation to parental sensitivity in the context of the wider literature is now considered in turn. Associations between parental sensitivity and constructs related to

children's internal representations are understudied. To enhance our understanding of the findings, I consider a landscape of research that broadly considers the influences of parenting on children's internal representations, rather than exclusively focusing on sensitivity and/or the precise dimensions of children's narratives considered in the current study.

5.5.1.1.1 Parental sensitivity and narrative coherence.

Higher parental sensitivity was associated with more coherent narratives in the children's story responses two years later. Considering this finding through the lens of attachment theory offers one way to consider possible mechanisms underlying the relationship. Theorists have proposed that a sensitive caregiving style is associated with children's attachment security, a proposition that has been supported by meta-analytic evidence (e.g., De Wolff & van IJzendoorn, 1997). Under this framework, the quality of early caregiving experiences is said to feed into children's internal working models, where they form expectations of themselves and others (Moss et al., 2009). Linking this to the children's responses to the story stems, previous studies have reported associations between early assessments of children's attachment and their later narrative coherence (Moss et al., 2009; Sher-Censor & Oppenheim, 2004; Wong et al., 2011). Indeed, it has been proposed that narrative coherence, more than the content of their narratives, may be most pertinent in understanding the structure and organisation of children's internal working models and attachment representations (Sher-Censor & Oppenheim, 2004; Yuval-Adler & Oppenheim, 2015). This also interlinks with the theory underpinning the Adult Attachment Interview (Main, Goldwyn, & Hesse, 2002) in which it is proposed that the overall coherence with which adults reflect upon and evaluate details of their early experience with key attachment figures can provide insight into their state of mind with respect to attachment (Fonagy, Steele, & Steele, 1991; Reese, 2008). Parental sensitivity itself has been less well-studied in the story stem literature and mostly considered in the context of children's representations of parental figures in their stories (e.g., Goodman, Aber, Berlin, & Brooks-Gunn, 1998), rather than narrative coherence. The current findings therefore make a unique contribution to a scant evidence base by indicating that the quality of sensitive interactions experienced in children's earliest years may be associated with their later abilities to construct and organise narratives in a coherent way.

5.5.1.1.2 Parental sensitivity and behavioural dysregulation.

Higher parental sensitivity was associated with lower behavioural dysregulation two years later, measured by assessing children's behaviours related to compliance and inhibition during the story stem task. It may be that children with more supportive parents can better rely on their caregivers as a source of observation, active guidance, and encouragement as they learn about the expectations and demands of their social environment. This reliable support could allow the child to more successfully internalise societal norms and values to facilitate more effective self-regulation. Partial support for the role of early parent-child relationships as a first model for patterns of self-regulation comes from a meta-analysis of 41 studies which revealed that positive (e.g., teaching, encouraging, guiding) and negative (e.g., criticism, harshness) parental control were both associated with children's regulation but responsiveness (e.g., synchronicity, warmth, positive affect) was not (Karreman et al., 2006). It could also be that being in receipt of more positive caregiving behaviours reduces negative arousal in the child, allowing them to more easily and flexibly adopt appropriate strategies for their self-regulation. Experiencing sensitive caregiving may promote children's construction of their internal cognitive frameworks, allowing for the clear organisation of expectations and behavioural strategies for them to draw upon when attempting to regulate their behaviour (Hawkins & Haskett, 2014; Macfie & Swan, 2009). Support for this link to children's internal representations comes from Birmingham, Bub, and Vaughn (2017) who reported that parental sensitivity predicted self-regulation and that this relationship was mediated by children's attachment history. Another possibility considers the bidirectionality of this relationship, in that children who display better behavioural regulation elicit more positive parenting practices.

5.5.1.1.3 Parental sensitivity and children's expression of disruptive and prosocial themes.

Early parental sensitivity was not found to be associated with children's expression of disruptive or prosocial narrative themes. This is an understudied area particularly in preschool populations and, to my knowledge, this is the first study to explore these specific relationships in an at-risk population of three- and four-year-olds. Drawing on wider literature though, the findings contrast with those of Laible et al. (2004; $N = 63$) who reported in a low-risk sample that warm parenting predicted expressions of prosocial story themes while harsh parenting predicted aggressive thematic content. However, although there is conceptual cross-over between the definitions of parental sensitivity, warmth, and harshness, differences in their

effects on children's functioning have also been noted (Mesman & Emmen, 2013). A study from Bovenschen et al. (2016) revealed no associations between children's narrative representations and foster parents' sensitivity in a small (N = 49) sample of three- to eight-year-olds. However, interestingly, further analysis of this same sample revealed associations between aspects of male (but not female) foster children's narrative representations (including aggressive content) and their foster parents' state of mind regarding attachment (Nowacki et al., 2015). Although the very small sample means that results should be interpreted with caution, the authors suggest that the findings indicate that caregiver's attachment representations, over caregiver behaviour, may be more salient for children's narrative representations. This suggestion is supported by previous evidence of correspondence between parent and child representations of attachment (Miljkovitch, Pierrehumbert, Bretherton, & Halfon, 2004; Pederson, Gleason, Moran, & Bento, 1998; van IJzendoorn, 1995). While it has been suggested that caregivers express their attachment representations through more or less responsive behaviour towards their children, meta-analyses of these relationships (e.g., Goldsmith & Alansky, 1987) reveal that the largest part of the influence of parental state of mind on their child's attachment security operates through mechanisms other than parental sensitivity (van IJzendoorn, 1995). Further study is needed to address the significant gap in our knowledge of how attachment representations are transmitted.

5.5.1.2 Aim 2: To explore whether intervention effects on narrative coherence (as measured through the story stems) were mediated by parental sensitivity.

Reflecting the full sample examined in the previous chapter, analysis of this subsample revealed that children whose families received the VIPP-SD intervention exhibited higher narrative coherence than children in the usual care arm (d = .32). It was hypothesised that improvements in parental sensitivity may have played a mediating role in the pathway between the intervention and children's narrative coherence. However, no evidence of such an indirect effect was found. There are several possible explanations for such a finding but here I consider a few alternative factors, other than improvements in parental sensitivity, which may have contributed to the group differences in narrative coherence.

Previous research has revealed close links between children's attachment security and their narrative coherence in response to story stems (Moss et al., 2009; Sher-Censor & Oppenheim, 2004). Given that the principal targets for the VIPP-SD programme were children's primary caregivers, one possibility is that group differences in narrative coherence may have been mediated by children's attachment representations post-intervention. Previous

research exploring the impacts of the VIPP-SD intervention indicates that it has the capabilities to improve attachment security, with meta-analytic evidence reporting an effect size of $d = 0.36$ (Juffer, Bakermans-Kranenburg, et al., 2017b). Although parental sensitivity and children's attachment security are linked, I have already discussed how sensitivity is just one of the multiple child, familial, and contextual factors considered to contribute to the development and maintenance of children's attachment representations (De Wolff & van IJzendoorn, 1997). Thus, parental sensitivity not appearing to have a mediating effect in this study does not rule out attachment history as a possible mediator.

Another possibility is that the VIPP-SD intervention may have promoted a different parenting behaviour, or set of parenting behaviours, other than parental sensitivity that supported children's expression of coherent narrative representations. For instance, improvements in communication style and empathy exhibited by the parents may have promoted skills related to coherence, as was the case in Reese et al.'s (2020) study of a parenting intervention. Or, if we continue to assume that narrative coherence is in some way linked to children's quality of attachment, then we could consider the possibility that the intervention promoted other domains of parental behaviours that have been previously linked to attachment security. For example, it may have promoted synchronicity, emotional support, or positive attitude, as suggested by De Wolff and van IJzendoorn (1997), or parental reflective functioning or mind-mindedness, as suggested by O'Hara et al. (2019) in their systematic review of video-feedback interventions. Each of these behaviours have only been found to be modestly linked to attachment security though and therefore the contributions of unique influences on attachment, that appear to be widely unaccounted for in studies, must be acknowledged. Given that a multidimensional approach to exploring influences on attachment has been recommended, perhaps it is also necessary to undertake such an approach when considering influences on narrative coherence, given its close theoretical links to attachment representations.

5.5.1.3 Aim 3: To explore whether intervention effects on children's behaviour (as measured through the CBCL) were mediated by parental sensitivity.

It was hypothesised that children whose families received the VIPP-SD intervention would exhibit lower levels of behaviour problems post-intervention and that this effect would be mediated by improvements in parental sensitivity. Although mean differences in behaviour problems were found between groups, no statistically significant mediating effect of

sensitivity was found. There are numerous factors that could be considered when interpreting these findings but here I focus on three specific possibilities.

Improvements in parental sensitivity post-intervention were more modest than has been seen in previous studies of the VIPP-SD intervention. For each subsample available in this analysis (n = 154 and n = 178), the group difference in parental sensitivity at the post-intervention follow-up had effect sizes of d = .11 and d = .23 respectively, compared to the combined effect size reported in a meta-analysis of previous VIPP-SD studies of d = .47 (Juffer, Bakermans-Kranenburg, et al., 2017b). The meta-analysis reported gains to be greatest when samples were recruited based on parental or environmental risk, including samples comprised of insensitive parents (d = .78), families living in poverty (d = .70), and adoptive families (d = .64) and lower effects when inclusion in the study was dependent on child-related risk including children recruited based on externalising difficulties (d = .34), as was the case in the current study, and children with autism (ds .24-.38). One possible explanation for the more modest effects of the current study may be the pragmatic nature of the trial's conduct, with 40 health professionals (predominantly health visitors and nursery nurses) delivering the intervention as part of a wider caseload in public health services. This approach contrasts with the tightly controlled research settings in which other trials have been conducted. A closer examination of the 12 studies included in Juffer et al.'s (2017b) meta-analysis reveals that of all them were delivered within research contexts, with most therapists being graduate-level research students, and only a minority reporting clinic-based professions (e.g., nursing, social work). Under research conditions, therapists are likely to have had greater capacity to deliver the intervention as prescribed. However, while the pragmatic nature of the current trial may have had some impact on weakening the effect of the intervention, it should also be noted that fidelity to the manual for the full HSHS sample was found to be high (94% of sessions met the minimum fidelity threshold, as reported in O'Farrelly, Barker, et al., 2021).

Another consideration is that the effects of the intervention may have been more pronounced for parents and children with a higher susceptibility to environmental enrichment. Differential susceptibility theory proposes that certain temperamental, physiological, and genetic characteristics act as 'plasticity factors', contributing to some individuals not only being more susceptible to negative outcomes in adverse environments, but also to the positive outcomes of supportive environmental conditions in a 'for-better-and-for-worse' manner (Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2007; Belsky & Pluess, 2009; Ellis,

Boyce, Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2011). A particularly compelling extension to this theory comes from Keers et al. (2016) who suggest that environmental responsivity is a complex, 'polygenic' trait, involving the accumulation of multiple plasticity factors of small effect. It has been proposed that by aggregating across more susceptible and less susceptible individuals, the effectiveness of parenting interventions may be underestimated, or even go undetected as a main effect, if such factors are not included as potential moderators. Such differential effects of the VIPP-SD programme have been reported for children with highly reactive temperaments (Klein Velderman et al., 2006), and the DRD4 7-repeat allele (Bakermans-Kranenburg, van IJzendoorn, Pijlman, Mesman, & Juffer, 2008), but not for highly reactive parents in the recent trial from Euser et al. (2021). It has been suggested that better understanding variation in the susceptibility of both parents and children to the potential enrichment of interventions could allow for the adaptation of programmes in a way that provides optimal "susceptibility fit" for families (Juffer, Bakermans-Kranenburg, et al., 2017b). For the present study, it may be that the programme had stronger effects on children and/or parents who are more sensitive and responsive to their environments but that these effects are hidden in the current analysis of the group as a whole.

Another possible explanation for the lack of mediating effect of parental sensitivity could be that the programme promoted other aspects of parenting that are important for children's behavioural development but were unmeasured in the current study. In a recent trial of the VIPP-SD programme with parents of preschool-aged twins, Euser et al. (2021) reported that the intervention did not enhance parental sensitivity but did have a positive impact on the level of caregivers' positive limit-setting. As well as targeting parental sensitivity, a key aim of the VIPP-SD programme is to promote a sensitive, consistent, but non-harsh approach to limit setting and discipline. This element of the intervention focuses on encouraging caregivers to reinforce children's positive behaviours, rather than engaging in "coercive cycles" in response to negative behaviour, and to provide clear explanations for times the child is prohibited from doing something. Such effects may be present but are unmeasured in the current study. This may be a worthwhile future avenue of exploration given that discipline style has been linked to children's behaviour problems (Mackenbach et al., 2014; Patterson, 1982).

Finally, although the present magnitude of effect on parental sensitivity is smaller than that found in Juffer et al.'s (2017b) meta-analysis, a consideration of the possible implications of an effect size of this magnitude should still be considered. Using conventional Cohen

criteria (0.2 = small, 0.5 = medium, 0.8 = large; Cohen, 1969), an effect size of $d = 0.11$ and $d = .23$ would be regarded as very small and small. However, there have been recent discussions around whether effect size benchmarks should be lower for social sciences research, with calls for studies aiming to explore children's developmental and educational outcomes to consider strength of association in accordance with domain-specific benchmarks of effect (Kraft, 2020; Stanley, Carter, & Doucouliagos, 2018). Such considerations may be particularly important for preventive interventions, whereby the transfer of small effects to large populations could be especially valuable, both for families and wider society (Euser et al., 2021). Based on the distribution of 1,942 effect sizes from 747 RCTs of interventions targeted at kindergarten- to high school-aged children, Kraft (2020) proposed the following baseline benchmarks for effects: less that 0.05 is small, 0.05 to less than 0.20 is medium, and 0.20 or greater is large. Therefore, while the smaller than anticipated effect found on parental sensitivity in the current study may partially explain the lack of evidence of mediation, the possible implications of an effect of this magnitude should not be dismissed entirely.

5.5.2 Strengths and limitations of the study

Here I discuss some of the strengths and limitations of the present study by considering elements of the study design, sample, and measurement of key variables in turn.

5.5.2.1 Study design

The longitudinal, pragmatic, experimental design of this study offers several important strengths. In their meta-analysis of video-feedback parenting interventions, O'Hara et al. (2019) emphasise the need for more studies to consider longer term follow-up outcomes (over 12 months post-intervention). The current study benefits from the inclusion of the child-reported story stem measure at the 24-month follow-up assessment, approximately 18 months after the VIPP-SD programme ended for families randomised to receive the intervention. Further to this, including measures across three time points over two years allows us to explore some of the developmental pathways of the constructs of interest. Also, the multi-site, pragmatic nature of the study provides important insights into the real-world application and effects of the VIPP-SD intervention when delivered as part of routine care, maximising the external validity of findings. Typically, parenting interventions are developed and initially tested in highly controlled, 'ideal' research conditions which can be helpful in the explanatory stage of intervention development but tell us little about the real-world feasibility or impacts of such a programme.

5.5.2.2 Study sample

Here, I consider two sociodemographic characteristics of this sample that may limit the generalisability of study findings: caregiver status and caregiver educational attainment.

First, this sample was almost exclusively comprised of mother-child dyads, with only a small proportion of families identifying the child's father as their primary caregiver (n = 10; 5%). This disparity in caregiver representation is not uncommon in the field, with O'Hara et al.'s (2019) study of video-feedback interventions noting a scarcity of studies that included any fathers and the small proportions of dads (<10%) in studies where they did participate. There has been a significant increase in father involvement in child caregiving over the last few decades and a large majority of children have close interactions and attachments with their fathers and other important caregivers (Bakermans-Kranenburg, Lotz, Alyousefi-van Dijk, & van IJzendoorn, 2019; Cabrera et al., 2018). The current study's findings do not account for the important contributions of paternal sensitivity on children's developmental outcomes, as has been previously established through meta-analytic evidence (Rodrigues et al., 2021). Further to this, despite the fact 81% of participating children were living in two-caregiver households, the dyadic nature of the study fails to consider the family processes underpinning the coparenting and triadic relationships, which have been found to be important influences in children's development (Teubert & Pinquart, 2010). Of particular relevance to this study, associations between aspects of triadic relational processes and children's outcomes appear to extend to children's representational worlds. von Klitzing and Bürgin (2005) reported positive longitudinal associations between parental triadic capacities and children's narrative coherence and expression of positive story themes, as well as their externalising behaviour problems.

Second, this sample was more highly educated than the general population (60% vs 41% of 25-34 year olds educated to graduate level in the UK; Census, 2011b). Efforts were made to avoid this commonly seen issue in research by primarily recruiting at developmental reviews conducted through the universal services provided by health visitors. These reviews are offered to all families when their children are 9-12 months and again at 2-2.5 years of age and have a national uptake of 77-90% (Public Health England, 2019). However, despite this approach ensuring we reached a more representative population of families in the screening phase of the study (49% educated to graduate level), this representation dipped for recruitment into the main phase of the trial. This suggests that future studies may need to undertake different or enhanced recruitment and retention strategies to ensure that research

seems accessible to all families. This could help ensure that research findings are as representative of the general population as possible. Breaking down such barriers is especially important in the field of developmental psychopathology, given that we know families with lower SES and educational attainment are less likely to engage with research studies but more likely to be in need of or seeking support for their child's behaviour (e.g., Heinrichs, Bertram, Kuschel, & Hahlweg, 2005).

5.5.2.3 Measurement of study variables

5.5.2.3.1 Measurement of parental sensitivity

A key component of the current study is the measurement of parental sensitivity. Here I reflect on the strengths of this methodological approach, as well as some considerations around the representativeness of the behaviours elicited through the parent-child interaction used, and the time-efficiency and cost-(in)effectiveness of such measurement.

Observational methodology. Undertaking direct observations of parent-child dyads using well-operationalised scales may have provided a more objective and valid picture of parental behaviours compared to measures that rely on individual's interpretation of items and a parent's capacity for self-reflection. Such self-assessment may be particularly inadvisable for measuring sensitivity given that caregivers can only report on child signals they notice, rather than those they miss or misinterpret (Rodrigues et al., 2021). Caregivers' ability to notice, correctly interpret, and appropriately respond to their child's signals is a key consideration in the assessment of their sensitivity. Further to this, observing parent-child interactions may be particularly useful when assessing interventions, as it has been suggested that caregivers may overestimate changes in their own behaviour following intervention (Patterson, 1982).

Home-based setting. Assessing parenting behaviours in the family home may have provided a much closer approximation of typical dyadic interactions, compared to observations captured in unfamiliar laboratory or clinic-based settings. Previous research suggests that home-based data collection may be particularly advantageous when studying children with behavioural difficulties as they may be less likely to inhibit their typical behaviours in a familiar setting (Aspland & Gardner, 2003; Webster-Stratton, 1985). This, in turn, may provide a more authentic and ecologically valid representation of parenting quality in response to more characteristic child behaviour. However, observing dyads in the home required researchers to remain in the same room as the family and use handheld video cameras to record them. This method of data collection is more obtrusive than the use of

remote-controlled cameras available in laboratory-based studies. It is possible that such close observation may have resulted in behavioural changes, although studies of such effects have reported minimal participant video reactivity in parent-child assessments (Antal et al., 2015; Bennetts et al., 2017; Semeniuk & Riesch, 2011).

Nature of the parent-child interaction. It is important to consider the degree to which behaviours captured during observation are representative of those that typically occur within the parent-child dyad. In the current study, parenting was observed during a low-stress, free play interaction. Play is a familiar family activity and the lack of specific task instructions may have enhanced the validity of the behaviours captured, as was found in Gardner (2000). However, play interactions only capture one component of family life. Caregivers spend a significant portion of their time completing caregiving tasks (e.g., bath time, mealtimes, outfit changes) during their daily interactions with their child. While some studies have reported a moderate to strong stability in sensitivity across contexts (e.g., Braungart-Rieker et al., 2014), parental sensitivity was found to be higher in naturalistic caregiving activities than in free play interactions in a study by Branger, Emmen, Woudstra, Alink, and Mesman (2019). Indeed, Ainsworth's original scales of maternal sensitivity were derived from observations of sensitivity to attachment behaviours in naturalistic, home-based, caregiving interactions. Further to this, the Erickson scales have most frequently been applied to higher-stress parent-child interactions with an element of problem solving or teaching in them (e.g., completing a puzzle that is too difficult for the child, constructing an Etch-a-Sketch image; Mesman & Emmen, 2013). It may be that low-stress, play-based interactions elicit a different set of parental responses that may be better suited to an altered assessment of sensitivity. Interestingly, Tamis-LeMonda, Bornstein, Baumwell, and Melstein Damast (1996) reported that parental responsiveness to language predicted children's language and parental responsiveness to play predicted children's play, indicating that not only can contexts elicit different parenting behaviours, but that these contexts may differentially inform our understanding of specific child outcomes. Overall, while the observational data collected may well be representative of a portion of daily parent-child interactions, we cannot assume it provides us with the complete picture.

Time-efficiency and cost-effectiveness. A key limitation to the measure of parental sensitivity was the time and cost required to collect and code the data. With participants living across six NHS sites, total researcher travel time typically ranged from two to five hours per family. In-person data collection brings several advantages. However, the increasing

accessibility of technology and internet for families means that developing parent-child interaction measures that can be administered remotely may become a more appealing alternative. For instance, Oliver and Pike's (2021) *Etch-a-Sketch Online* (ESO) observation tool can be delivered over video call and has demonstrated high inter-rater reliability and convergent and predictive validity. Training in the use of the parental sensitivity scales to an acceptable standard was also expensive and time-consuming (eight months between the initial training course and the completion of two reliability sets), as was the coding of data (45-60 minutes per case; ~360 hours of coding for 180 families across two timepoints). The high resource needed for this coding meant I was the only researcher available to code the interactions. This limited the study sample size and possibly introduced issues related to researcher drift and unreliability, given that time between reaching reliability and dataset completion spanned an 18-month period. In recent years, there has been a call for the development of more time-efficient and cost-effective observational measures of parent behaviours to bridge the research-to-practice gap. One example of this can be seen with the Atypical Maternal Behaviour Instrument for Assessment and Classification, a measure of disrupted caregiving (AMBIANCE, Bronfman, Madigan, & Lyons-Ruth, 2009). The development of the AMBIANCE-Brief cut administration time from ~60 minutes to ~6 minutes. It was also found to be feasible for delivery in community settings, with good convergent and concurrent validity (Cooke, Eirich, Racine, Lyons-Ruth, & Madigan, 2020; Haltigan et al., 2019; Madigan et al., 2021). Such alternatives may lead to observational assessment becoming more feasible in clinical settings, scaled-up intervention programmes, and studies across diverse and low-resourced contexts. Investment in developing such measures may be particularly pertinent for the coming years where we are likely to see cuts to research funding as a result of the hardships brought by the COVID-19 pandemic.

5.5.2.3.2 Measurement of children's representations

The previous chapters in this thesis have already evaluated the use of the doll's house play story stem assessment battery in detail (e.g., Chapter 3, section 3.5.2.2.1). The preceding studies indicated that it is a feasible measure with good inter-rater reliability, and evidence of concurrent validity with related constructs on parent- and teacher-rated outcomes. What the findings of the present study continue to demonstrate is that by including a measure of child-reported outcomes, we are accessing information that we do not obtain through adult-rated measures alone. The findings highlight the enhanced understanding bringing children's perspectives into standard research practice can offer. Measuring constructs using a play-

based measure with children as young as three years old meant we gained unique insight into the ways in which parental sensitivity predicted later child functioning. This strengthens the case already made that involving children in research as another source of information provides valuable insight on top of what the adults around them offer.

5.5.3 Conclusion and implications

The findings of this study suggest that the quality of early caregiving children receive may be an important contributing factor in the development of children's narrative coherence and behavioural regulation. These findings add to our understanding of the unique and consequential role caregivers play in their children's lives. They also highlight the heightened insight that can come from including measures of children's internal representations (in this case to measure narrative coherence) and direct observations of child behaviour (in this case to measure behavioural dysregulation) in investigations of the developmental influences of caregivers on children's functioning.

The study also suggests that a home-based video-feedback parenting programme targeting parental sensitivity and sensitive discipline style can promote children's behavioural development and narrative coherence. These findings are important as they highlight the possible gains for children and families following early parenting supports. Children exhibiting early signs of behavioural difficulties are at an increased likelihood of long-term psychological, educational, and societal difficulties. The current research findings promote the importance of providing parents and families with effective, targeted parenting supports at the earliest signs of difficulties. This programme may offer us one component to be used in a wider evidence-based package of care for families, forming a powerful pathway of prevention and intervention through the family life course.

However, we are still lacking in a complete understanding of the treatment mechanisms underlying the relationships found. Further study of the HSHS sample itself, as well as future studies, may benefit from undertaking a multidimensional approach to exploring the effects of parenting on children's behavioural outcomes. Possible additional parenting behaviours that might contribute to the relationships seen in this data, and that VIPP-SD has previously been shown to promote, include a sensitive discipline style (Euser et al., 2021; Pereira et al., 2014), family cohesion (Negrão, Pereira, Soares, & Mesman, 2014), parental self-efficacy (Poslawsky et al., 2015), conflict management (Woolley et al., 2007), and attitude about caregiving (Groeneveld et al., 2011; Werner, Vermeer, Linting, & van IJzendoorn, 2018). Further to this,

future research may benefit from adopting a differential susceptibility approach in analysis to identify whether particular parent- or child-related factors contribute to families being more or less susceptible to the environmental enrichment provided by effective interventions.

To conclude, this study enhances our understanding of the importance of early caregiving experiences on aspects of children's behavioural development. Nonetheless, the mixed and modest effects reinforce the notion that children's development is influenced by a diverse range of multi-layered child, family, sociocultural, and contextual factors. Importantly, the relatively novel use of a doll's play narrative-based measure with three- and four-year-old participants in this study allows us to better understand the ways in which external experiences may shape children's internal functioning. Integrating such measures, which are brief, flexible, and portable, more widely into research practice can not only allow us to better understand the impacts of parenting interventions on children's external behaviours but also the internal systems that guide and shape their behaviour.

CHAPTER 6: Discussion

6.1 Chapter rationale

The central aims of this thesis were to explore the integration of children's perspectives into investigations of the presentation, development, and intervention of early behavioural difficulties in very young children and to investigate the influences of the parent-child relationship on children's behavioural development. The studies within this thesis are characterised by three overarching themes: behaviour problems in young children, the role of parenting and the parent-child relationship in children's earliest years of development, and the integration of very young children's (under five years old) perspectives into our understanding of these trajectories.

In this chapter, I consider a set of overarching findings that can be drawn out from the previous chapters. I offer a discussion of some of the strengths and limitations of the studies and the overall approach taken in this thesis before reflecting upon some of the possible implications for research and clinical practice.

6.2 Chapter overview

In this final chapter I draw together and discuss a core set of findings derived from the thesis as a whole. I then consider notable strengths and limitations of the approaches undertaken across the three studies and reflect on some of the possible implications of this work for research and practice. Figure 18 provides a summary of the key components of this chapter.

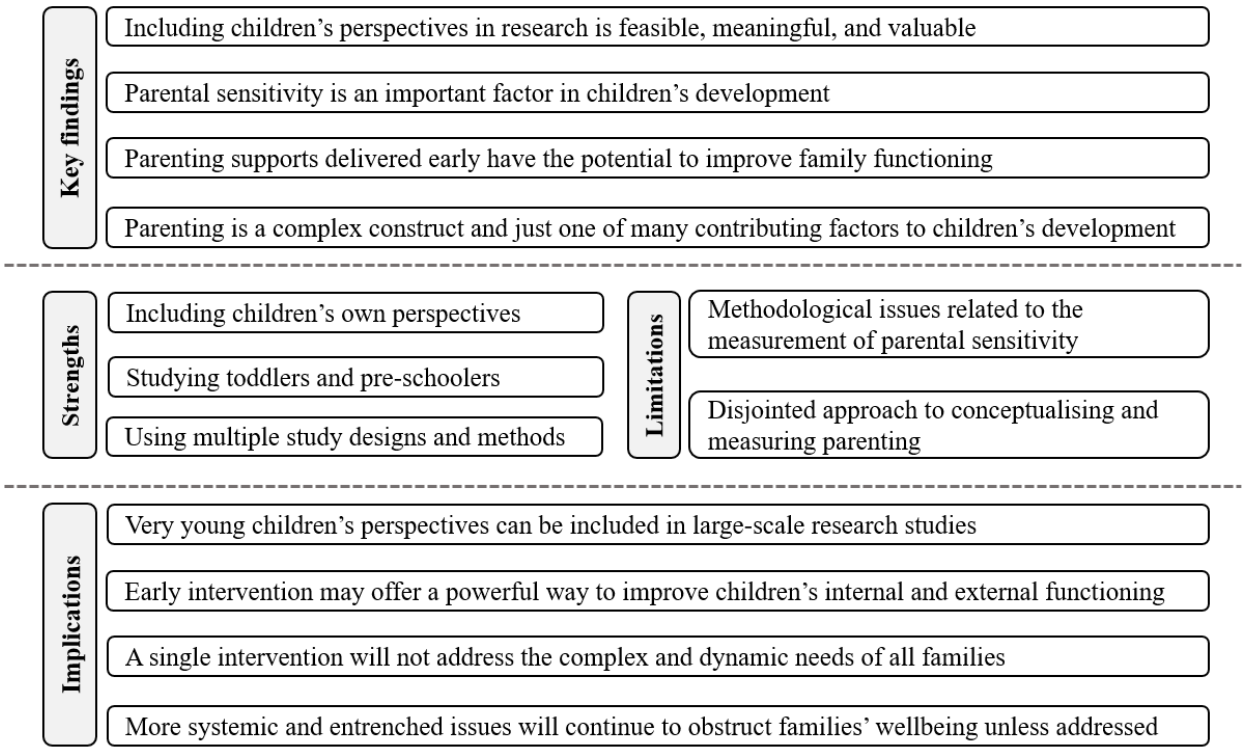


Figure 18. Summary of the key components of the discussion chapter.

6.3 Summary of findings

The overarching aim of this thesis was to explore the integration of children's perspectives into investigations of the presentation, development, and intervention of early behavioural difficulties in very young children. It also aimed to investigate the influences of parental discipline, parental sensitivity, and a video-feedback parenting programme on children's outcomes. A comprehensive discussion of the aims and findings of each of the studies within this thesis have been provided in their respective chapters. Here, through an appraisal of the thesis as a whole, I have drawn out four overarching findings that I consider encompass core components of the study results. Figure 19 depicts the ways in which these broad elements relate to some of the specific results of each chapter.

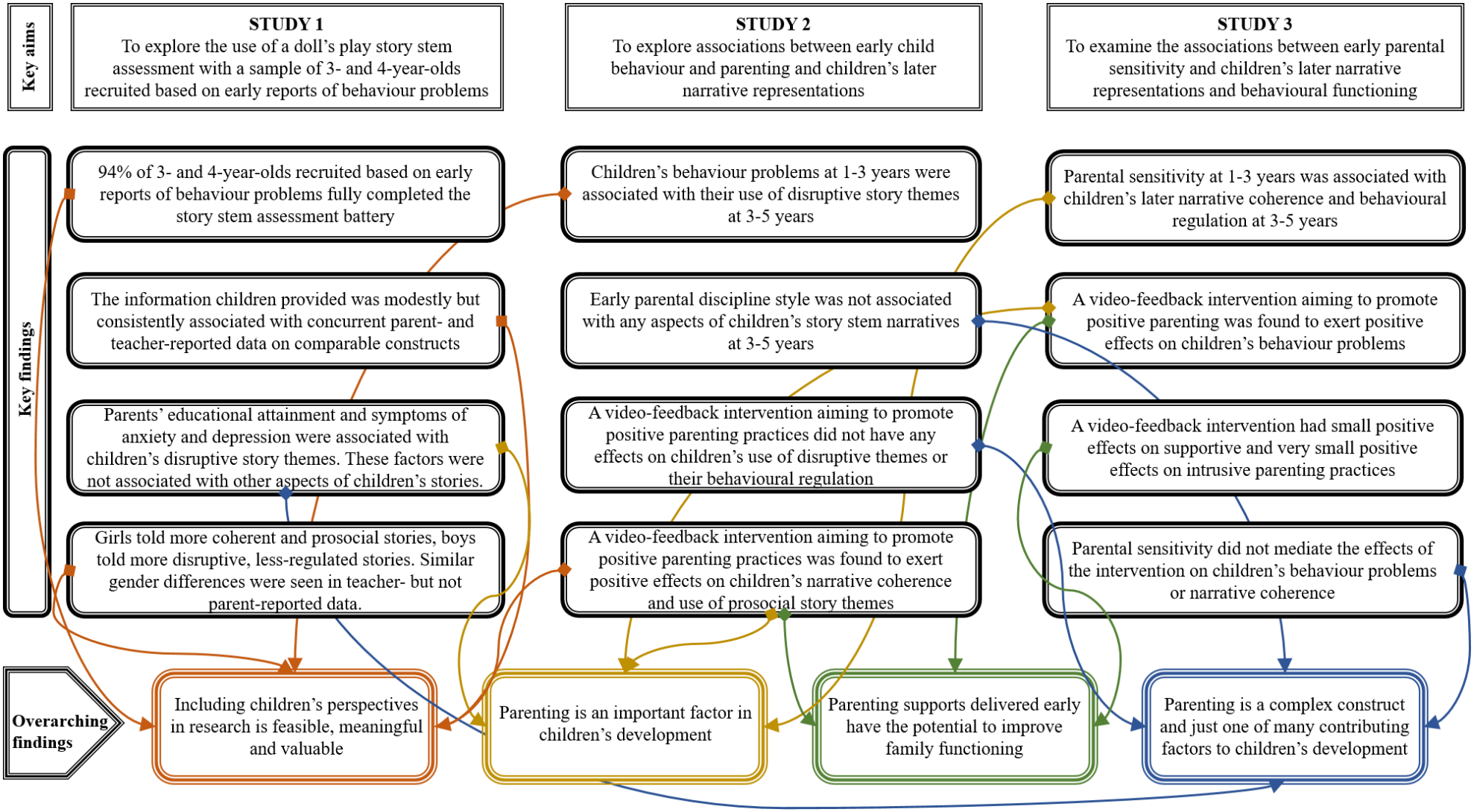


Figure 19. A depiction of the ways in which key findings of each study's chapter feeds into the overarching findings summarised in this section.

6.3.1 Including children's perspectives is feasible, meaningful, and valuable

Children have the right to be involved in matters that affect their lives – a right that is largely overlooked in developmental psychology research. The conceptualisation of this research was rooted in the idea that, when given the right tools and opportunities to do so, even young children can offer insight that enhances our understanding of their development. This thesis provides support for the feasibility and value of including very young children's perspectives in studies about their behaviour and development.

Chapter 3 studied the use of a doll's house play story stem assessment battery in a sample of three- and four-year-olds recruited based on early parental reports of behaviour problems. The assessment battery was feasible to deliver and acceptable to families, with 94% of children fully completing the measure. This research demonstrates that play-based measures are an engaging way to facilitate children's involvement in research, even for populations considered to be more likely to struggle with their attention and behaviour. However, assent rates were lower in the small subset of children ($N = 11$) who had received or were undergoing assessment for a clinical diagnosis of ADHD or ASD (diagnosis: 45% fully completed; no diagnosis: 96% fully completed). Further studies focusing on the use of child-reported measures with these groups is warranted to explore whether additional adaptations to the measure's design and administration may better facilitate children's involvement. At present, this is a much understudied area with only a handful of studies utilising narrative representation tasks in (small, older) samples of children with diagnoses of ADHD (e.g., Scholtens et al., 2014; 6-10 year olds; $N = 89$) and ASD (e.g., McCullough, Stedmon, & Dallos, 2014; 9-13 year olds; $N = 4$).

Concurrent correlational analyses revealed a modest degree of convergence between constructs derived from the children's responses to the story stem battery and related domains of parent- and teacher-reported child behaviour. The magnitude of these associations was small (rs .14-.26) but consistent with previous studies of multi-informant psychological research (e.g., Achenbach et al., 1987; McConaughy et al., 1992). These findings suggest that the perspectives of three- and four-year-olds correspond with adults' perceptions of their behaviour to some degree but, at the same time, offer us unique insight into aspects of their development. The contextual and situational specificity of children's behaviour may partly explain differences in reports from parents, teachers, and children themselves. For instance, behavioural and social expectations vary, as do the adults enforcing these expectations, across home and educational settings. These differences may result in divergence in children's

ability and willingness to comply with rules and expectations across settings and, in turn, the cross-context reports of child behaviour (Smith-Donald et al., 2007). Teachers are more likely to see children during task-oriented and social situations and have other children in the same situation as a comparison to inform their judgement (de Nijs et al., 2004). Conversely, parents are often restricted to comparisons to siblings, or their perceptions of the behaviour of a small number of peers, when judging the significance of a behaviour. Children themselves are aware of their behaviour across contexts and therefore child-reported data may partially reflect their behaviour and experiences both in the home and at school (Arseneault et al., 2005). Given that narrative-based methods aim to elicit aspects of children's internal representations, children may be providing us with access to aspects of their internal systems that the adults around them are unable to report on. This suggests that multi-informant approaches may provide us with a more holistic and thus valid picture of children's development.

The findings of Chapters 4 and 5 highlighted the unique value that measures of children's perspectives can bring. In Chapter 4, children's behaviour problems measured at one and two years old predicted their later expression of disruptive story themes, and positive effects of a parenting intervention on children's narrative coherence and expression of prosocial themes were revealed. In Chapter 5, early parental sensitivity was found to be associated with children's later narrative coherence and behavioural dysregulation. Understanding of the early influences of child behaviour and parenting practices on constructs thought to be related to children's internal representations would not have been gained through adult-reported data alone.

Together, these findings suggest that the use of developmentally sensitive, play-based tools can allow preschool-aged children to provide unique and valuable insight into their development.

6.3.2 The parent-child relationship is an important factor in children's behavioural development

The important role of caregivers in their children's development was highlighted through several findings in this thesis. In Chapter 3, parents' educational attainment (r = -.13) and their symptoms of depression (r = .19) and anxiety (r = .17) were concurrently associated with children's expression of disruptive story themes. These associations were of a small magnitude but suggest that parental characteristics may be partially reflected in children's

(in)abilities to respond to issues in a non-disruptive way. More generally, this highlights the importance of parents having access to adequate support for feelings of low mood and anxiety, as this may also bring benefits to their children. In Chapter 5, parental sensitivity, measured when children were one and two years old, was associated with children's narrative coherence and behavioural dysregulation at three and four years old. This contributes to our understanding of the ways in which early parenting quality may relate to aspects of children's understanding, interpretation, and response to their worlds.

6.3.3 Parenting supports delivered in children's earliest years have the potential to improve family functioning

Findings in Chapters 4 and 5 suggested that an early intervention in the form of a brief, home-based, video-feedback parenting programme (VIPP-SD; Juffer et al., 2008b) had the potential to improve aspects of child and family functioning. In Chapter 4, group differences were found on children's narrative coherence and their expression of prosocial themes during the play-based story stem battery. In Chapter 5, group differences were found on children's behaviour difficulties. Together, these findings suggest that parenting interventions delivered in the first few years of life may not only impact upon children's external behaviours but also components of their internal systems, as measured through the story stem battery. However, the current set of studies did not provide clear insight into the mechanisms that might underlie such group differences. Chapter 5 provided a signal of an effect on caregivers' supportive parenting practices but this was a smaller-sized effect than anticipated based on existing evidence (Juffer, Bakermans-Kranenburg, et al., 2017b). Nonetheless, this provides an indication of the malleability of parenting, even when an intervention is delivered under pragmatic conditions. Extending this analysis out to the full sample of HSHS families could provide us with a more complete picture of the intervention effects on parenting style, as the current analysis only includes half of participating families. A closer look at alternative mediating or moderating factors, such as sensitive discipline style, may allow us to better understand how this intervention impacted upon children's external behaviour and a measure of their internal systems.

6.3.4 Parenting is a complex construct and just one of many contributing factors to children's development

While a small portion of the findings in this thesis pointed to the ways in which parenting can predict children's outcomes, the majority of findings indicated that parenting –

and specifically, parental discipline and parental sensitivity – are not the only important influences in children's functioning. In Chapter 3, parents' educational attainment and symptoms of depression and anxiety were not associated with any other aspects of children's stories besides disruptive themes. In Chapter 4, parental overreactive and lax discipline styles were not associated with any components of children's narrative representations. An intervention programme aiming to promote positive parenting practices did not have effects on children's use of disruptive themes or their behavioural dysregulation. In Chapter 5, early parental sensitivity did not predict children's later narrative coherence or prosocial themes. Collectively, these findings suggest that individual features of parenting do not operate in a deterministic manner. Other influences, including child- and parent-level individual factors (e.g., temperament; Belsky & van IJzendoorn, 2017), as well as wider environmental (e.g., household chaos; Coldwell, Pike, & Dunn, 2006) and contextual (e.g., social support, poverty; McConnell, Breitkreuz, & Savage, 2011; Steele et al., 2016) factors that were unmeasured in these studies, may have also contributed to the relationships found. For instance, while this thesis focused almost exclusively on the role of mothers in their children lives (96% of included caregivers), we know that 85% of participating children were growing up in two-caregiver homes, with a smaller proportion living in multigenerational households. The potentially significant contributions of wider family systems on children's outcomes are overlooked (Fagan, Day, Lamb, & Cabrera, 2014; Panter-Brick et al., 2014; Rodrigues et al., 2021). Overall, while there is a good theoretical basis and evidence base to suggest that parental discipline and sensitivity are important influences in children's development, what the current findings suggest is that they are just two of many factors that can promote children's development and that, when these practices are less optimal, there are other factors that can buffer or compensate against negative effects.

6.3.5 Aspects of the richness and nuance of family life and child development are absent in these findings

Before moving on to consider the strengths and limitations of this work, I wanted to end my discussion of the studies' results by reflecting on what might be missed when we focus exclusively on outcomes as a marker of our understanding into child development and family functioning. Exploring associations, effect sizes, and group differences offers a helpful way for us to conceptualise the magnitude of findings and make direct comparisons to other research studies. However, this focus on statistical results tells us little about what is meaningful to children and families themselves (O'Farrelly, 2021), nor what components of

interventions work for whom and under what circumstances (Bonell, Fletcher, Morton, Lorenc, & Moore, 2012). Process evaluations that use qualitative and quantitative methods offer us one way to better understand these nuances (Oakley, Strange, Bonell, Allen, & Stephenson, 2006). Such data can complement statistical outcomes, and aid our interpretation of findings, by exploring the implementation, receipt, and setting of interventions (Toroyan et al., 2004).

Having spent the last five years working with the families included in these studies, spending an estimated 800 hours in their homes, and a further 500 hours watching videos of their family moments shared together, it is odd to see the essence of the work boiled down to the interpretation of a collection of tables. Recently, I contacted Healthy Start, Happy Start families with our yearly update newsletter. This update came at least two years after the families were last visited as part of the trial, and four years since those randomised to the intervention received the last of their six programme visits. Despite this, several families responded to the update telling me how deeply helpful they had found the programme visits and about the positive impacts it had had on their relationship with their child. These replies came through as I was sat pondering the small effect size found in this set of studies relative to the effect size of the VIPP-SD meta-analysis (Juffer, Bakermans-Kranenburg, et al., 2017b). Of course, a few emails do not supersede statistical findings. However, they allowed me to reflect on the fact that so much of the richness and nuance of family experience is lost in the reporting of clinical trials. Thankfully, a process evaluation was undertaken as part of the wider HSHS study. As part of this, interviews were conducted with families and programme therapists about their experiences of receiving and delivering the intervention. Analysis of this qualitative data may help to embed the perspectives of parents and health professionals in our understanding of this programme. To build on the findings of this thesis and further integrate children's own perspectives, the children's story stem data could also be explored through a qualitative framework in an effort to capture a rich understanding of children's views of their early experiences.

6.4 Reflections on the strengths and limitations of this thesis

This section critically reflects upon three strengths of this research before considering two ways in which its implications may be limited by aspects of the methodological approach.

6.4.1 Notable strengths

6.4.1.1 Including children's own perspectives

Central to this thesis was the integration of a measure of children's own representations into a set of studies exploring associations between child behaviour, parenting, and early intervention. The elicitation of these insights from children aged only three and four years old is an important and unique strength of this work. Over three decades ago, the United Nations emphasised the need and value of including children's perspectives in research and policy starting in the earliest years (United Nations Committee on the Rights of the Child, 2005; United Nations Convention on the Rights of the Child, 1989). Despite this recognition, most work in the area continues to focus almost exclusively on school-aged children. Methodological approaches suited to older children are less well-matched to the attentional, cognitive, and linguistic developmental stages and abilities of younger populations. The use of the doll's house play measure in this thesis offers one developmentally sensitive way for the inclusion of children's voices. Its use promotes the importance and value of including children's own perspectives in early intervention research. However, this narrative-based, representational method offers just one level of understanding of children's experiences of their worlds. Future studies may benefit from mixed method approaches that cast the methodological net wide, providing us with different or complementary insights and understandings to those offered by representational methodologies. Supporting children to express themselves in a variety of forms has the potential to increase the depth, breadth, and value of their participation (Darbyshire, MacDougall, & Schiller, 2005). Further to this, future studies may extend children's participation to involvement in the data analysis through techniques such as respondent validation (e.g., Flewitt, 2005), or to roles in advisory groups during the study design phase (e.g., Emond, 2002), rather than conducting analysis exclusively through the lens of adult researchers.

6.4.1.2 Studying toddlers and pre-schoolers recruited based on parental reports of behaviour problems

A significant proportion of previous research in this area focuses on non-clinical, low-risk populations, or only includes populations at a stage when clinical difficulties have become embedded. This thesis focuses on a unique and understudied population of children, whereby the sample were followed from toddlerhood – when they were identified as being at an elevated likelihood of enduring behavioural difficulties based on parental report – to aged

three and four. Children typically begin to exhibit challenging behaviours from about 12 months onwards, with a peak at two years, and a decline between three and four years (Alink et al., 2006; Tremblay et al., 2004). Given these developmental trajectories, this thesis studies children at an especially sensitive time of their behavioural development, providing us with unique insight into the pathways and environmental influences of children's difficulties. In particular, the current research contributes to the relatively scant body of findings investigating the internal representations of preschool-aged children exhibiting early signs of behavioural difficulties. Studies of these populations are important. We know that the majority of young children who exhibit elevated behaviour problems in their earliest years do not go on to develop pervasive or persistent conduct disorders but, when these problems do endure, the consequences can be far-reaching (Fergusson et al., 2005; Gutman et al., 2018). Utilising unique methods to gain a more complete understanding of children's experiences, particularly during such a sensitive stage of behavioural development, could enhance our knowledge of the most appropriate and effective forms of support for children and families.

6.4.1.3 Study designs and methodologies

This thesis used multiple methodologies that differed in their detail and approach to investigate the research questions of interest.

In Chapter 3, the detailed coding of a doll's house play representational measure of children's narratives provided a unique understanding of children's development. Concurrent correlational analysis was used to establish if and how children's perspectives were related to those of their primary caregivers and teachers. As well as the novelty of the child-report measure, this multi-informant approach benefitted from the use of two psychometrically-robust questionnaire measures of child behaviour and a detailed semi-structured clinical interview with children's caregivers. This allowed for a more thorough consideration of the ways in which children's narrative representations related to their functioning. Particularly notable is the high response rate achieved from children's teachers (76%). This allowed for the inclusion of a second informant, who was possibly unaffected by the biases that can come with parental report (e.g., Ringoot et al., 2015), for a significant proportion of children.

Chapters 4 and 5 advanced on the findings of Chapter 3 by utilising longitudinal designs. Associations between early child behaviour and parenting practices, measured at only one and two years old, and children's later narrative representations and behavioural development at three and four years old were explored. This allowed children's development over time to be

examined, in the context of potentially risky or protective environmental exposures. Considering early behavioural and parental predictors of children's later internal systems, measured through the story stem assessment battery, is particularly unique as most of the previous research has explored these relationships concurrently. A key strength of the longitudinal designs was the high degree of retention achieved over the two years of data collection (85% for this subsample). It should be kept in mind though that the regression modelling conducted in these studies restricts our understanding to correlational, rather than causational, relationships. Although these statistical approaches were adequate to answer the research questions at hand, a more sophisticated approach to analysis may have enhanced the understanding attained. For instance, structural equation modelling could have allowed for a more complex consideration of the inter-connected influences of parenting. One option could have been to create a latent construct of 'positive parenting' from the variables related to parental behaviours (e.g., parental sensitivity, discipline) or characteristics (e.g., psychopathology). Such an approach gets closer to the multifaceted and intricate nature of parenting, rather than exploring the effects of singular variables. Learning these techniques, in parallel to collecting and coding the data, went beyond the resource of this research. With hindsight though, prioritising the attainment of these statistical skills would have allowed me to capitalise on the true richness of the data in a way linear regression modelling did not permit.

Chapters 4 and 5 also benefitted from the experimental design of the Healthy Start, Happy Start study. This design meant that the effects of a video-feedback programme (VIPP-SD; Juffer et al., 2008b) on children's narrative representations, behavioural functioning, and parental sensitivity could be explored. The inclusion of families who were partaking in a larger trial meant that the current studies are enhanced by the robust, rigorous, and unbiased procedures undertaken as part of a well-conducted randomised controlled trial. A particular strength of the delivery of VIPP-SD in this research is that it was delivered under pragmatic conditions. This allows us to understand the impacts of such interventions when delivered in real-world, routine health service contexts, as opposed to the tightly controlled, high-resource research settings such programmes are most typically designed and tested in. A further strength of the examination of intervention effects in Chapter 5 was the highly detailed measurement of changes in parental sensitivity. Using the Erickson scales required intensive training and the establishment of a good level of reliability. Undertaking observational approaches circumvents issues related to parents' self-reports, including biases in their recall

and interpretation of their own parenting behaviours, and the ways in which other factors, such as parents' mood or anxiety, may confound their reports. However, for the current thesis, this rigorous approach also brought limitations that are discussed further in the following section.

6.4.2 Notable limitations

Here, I focus on two specific limitations of the current research, both of which relate to the measurement and conceptualisation of parenting.

6.4.2.1 Methodological issues related to the measurement of parental sensitivity

While the advantages of measuring parental supportiveness and intrusiveness using the Erickson scales have been acknowledged, recognising the limitations of selecting such a highly detailed approach to measurement is also important. The process of undertaking the training and establishing reliability on the Ainsworth and Erickson scales was time-consuming and reliant on a fine-tuned, nuanced set of observational skills. Due to this, I was the only student who reached acceptable levels of reliability in the available timescale. Coding of the pre- and post-assessments of parent-child interactions for the participating families was therefore conducted by me alone. In my view, this has three important consequences for the current research and its potential implications.

First, it significantly reduced the sample size. I had met 193 (64%) participating HSHS families during data collection. In an effort to reduce the effects of bias, only families where I had not collected their baseline data were included in analysis. This decision halved the size of the sample, limiting our understanding of the potential effects of the VIPP-SD intervention on parents' sensitivity for the full HSHS cohort. Still, of those families retained in analysis, I had met 91 (49% of participants included in analysis) at one or both of their follow-up assessments. Potential effects of bias will have been reduced by the time that passed between collection and coding of data. However, I cannot be considered entirely independent as my awareness of participant characteristics may have still affected my coding.

Second, confidence in the quality of coding is reduced as there was no second coder to undertake checks for drift in the 18 months between establishing reliability and completing coding of the dataset. My understanding or interpretation of the coding scales may have shifted in this time. With hindsight, one way to check for drift as a solo coder could have been to rescore video interactions already coded and assess the degree of agreement between each set of scores. Issues with the quality of coding related to drift may be one explanatory factor

for the smaller than anticipated effects of the VIPP-SD programme on parental sensitivity, despite a robust existing evidence base and the intervention effects on children's functioning seen in the current set of studies.

Third, thinking more broadly about the implications of this limitation, it raises questions about the viability of undertaking such intensive approaches to assess parenting in large-scale research and clinical settings. The challenges of finding coders with the time and resource to undertake training, establish reliability, and complete coding for this thesis would likely only be heightened in standard research or clinical practice, particularly in low-resourced contexts. As detailed in Chapter 5, an especially valuable future direction for this area is the more widespread development and testing of tools that bring the benefits of observation but that are briefer and simpler to administer and establish reliability on (e.g., the development of the AMBIENCE-Brief; Madigan et al., 2021). Such tools may allow for the more frequent inclusion of observational assessments in scaled-up intervention programmes as well as studies in diverse and lower-resourced contexts (e.g., in low- and middle-income countries or time-poor clinical settings). Progress in this area may help to address the overwhelming imbalance of theoretical and empirical work that focuses almost exclusively on parenting in North America, Western Europe, and Australia. A review from Tomlinson, Bornstein, Marlow, and Swartz (2014) revealed that only 2.3% of 2002 to 2013 articles exploring infant mental health included data from low- and middle-income countries – where 90% of the world's infants are growing up (Haub & Kaneda, 2013). A review of child and adolescent mental health articles published between 2002 and 2011 revealed that, in those ten years, there was not a single publication from 42 countries (Kieling & Rohde, 2012). These 42 countries collectively housed 76 million of the world's children and adolescents. In keeping with this, Arnett (2008) reported that 96% of psychology research participants are from what has been termed WEIRD (western, educated, industrialised, rich, democratic; Henrich et al., 2010) countries. These voids in our global understanding have led to assumptions about the universality of the effects of 'westernised' parenting practices on child development, stunting the development of evidence-based interventions, policies and practice in low- and middle-income settings (Tomlinson et al., 2014).

6.4.2.2 Disjointed approach to the conceptualisation and measurement of parenting

There is a disjointed approach to measuring parenting in this thesis. Chapter 3 considers parental education and psychopathology, Chapter 4 parental discipline, and Chapter 5 parental sensitivity. While individually important, each are just one component of family life. Future

studies could benefit from undertaking a more multidimensional and integrated consideration of parenting, as well as a wider look at other factors affecting parents' capacity to parent in the way they would like to. Child development is a product of dynamic relationships between individual, familial, and environmental factors occurring together or over time. I will now discuss three specific factors that were unaccounted for in the current set of studies but that may have exerted main, mediating, or moderating effects on parenting and child outcomes.

First, temperamental traits may affect the parent-child relationship (Belsky & Fearon, 2008; Groh et al., 2017; van IJzendoorn & Bakermans-Kranenburg, 2012). Of particular interest to the current set of findings are studies suggesting that relationships between emotional reactivity, parenting quality, and children's behavioural development may operate in a manner consistent with differential susceptibility. This theory suggests that some children may be especially likely to benefit from the environmental enrichment successful parenting interventions provide (e.g., Bakermans-Kranenburg et al., 2008; Klein Velderman et al., 2006). Second, low parental self-efficacy may affect child outcomes directly but also indirectly via parents' capacity to engage in supportive and productive parenting practices (Albanese, Russo, & Geller, 2019). This factor may be particularly interesting to consider in the context of interventions. The positive, strengths-based focus of the VIPP-SD programme could well increase parents' belief in their ability to support and care for their child and alter their perception of their child's behaviour. Third, living in poverty can place acute stress on families in a way that other demographic factors (e.g., education level) do not fully capture. Such circumstances can influence parents' capacity to interact with their children in the way they might like to. A systematic review from Huaqing Qi and Kaiser (2003) revealed that preschool-aged children living in low-income families were significantly more likely to exhibit behaviour problems (31%) than the general population (3-6%). Future studies may benefit from including measures of pervasive hassles and stressors that oftentimes face families living in low-resourced households or communities (e.g., Perceived Stress Scale; Cohen, Kamarck, & Mermelstein, 1983; Parenting Stress Index; Abidin, Flens, & Austin, 2006) to aid our understanding of the ways these unique stressors may affect the relationships studied in this thesis.

While inclusion of measures related to these, or other, factors may have enhanced our understanding of the findings, collecting copious amounts of data in a single assessment risks fatiguing participants and compromising the validity of the data. One alternative to lengthy assessment batteries is the utilisation of synthesising methods such as systematic reviews,

meta-analyses, and umbrella meta-analyses that bring together the findings of multiple studies and increase statistical power for the detection of small effects.

6.5 Implications for research and practice

Here I consider four areas in which the current findings may help to inform research and practice.

6.5.1 Very young children's perspectives can be included in large-scale research

This set of studies indicates that including preschool-aged children's own perspectives in research is feasible and valuable, demonstrated by the successful use of a developmentally sensitive tool facilitating the participation of 256 three- and four-year-olds. When undertaking a multi-informant approach, future research studies can not only consider second caregivers, teachers, and clinicians, but also children themselves as an important complement to the information primary caregivers provide. Views on the specific purpose and use of multi-informant data varies (e.g., aggregating data across informants or analysing data separately with the view that each informant provides a valid but unique perspective; De Los Reyes, Thomas, Goodman, & Kundey, 2013), but key to including very young informants is using a mode of communication that is familiar and engaging. Play- and art-based methods may provide a particularly inclusive way to accommodate for variance in children's linguistic and cognitive abilities at this age (e.g., Blaisdell, Arnott, Wall, & Robinson, 2019; Janik Blaskova & Gibson, 2021; Koller & San Juan, 2015).

There are several learnings from this thesis that may be especially helpful for researchers and practitioners to keep in mind when considering the use of similar child-informant measures. First, the careful piloting of the story stem measure with 20 children across nursery and home settings proved to be a highly instructive step in its adaptation for use in the current set of studies. This was reliant on the establishment of relationships with community settings and local families, as well as a timeframe of approximately four weeks to conduct the piloting and iteratively adapt procedures. Researchers in future studies should ensure that they have allotted sufficient time and resource for similar piloting procedures. Second, the measure was administered by a relatively experienced research team who were well-versed in working flexibly, adapting to families' needs, engaging young children, and maintaining a continual awareness of fatigue or an unwillingness to participate in the children. Researchers undertook half a day's training on the measure and completed three practice cases. A greater degree of training may be required for less experienced teams. Finally, the measure was relatively time-

consuming to administer (15-25 minutes per child) and code (45-60 minutes per case). Such time and staff resource may not be available in other research studies and we also know less about the use of such measures in clinical settings.

6.5.2 Early intervention may offer a powerful way to improve children's internal and external functioning

The findings of this thesis provide support for the potential of early intervention and insight into the utility of VIPP-SD, a brief behaviourally-focused video-feedback programme.

The Early Intervention Foundation estimated that £16.6 billion is spent on 'late interventions' each year in England and Wales (Chowdry & Fitzsimons, 2016). Most of this cost was estimated to fall to local authorities (£6.4 billion) and the NHS (£3.7 billion). The potential savings that can come from effective early interventions is substantial (Doyle et al., 2009). Findings in this thesis indicated that an intervention delivered to the caregivers of one- and two-year-olds had positive effects on children's behaviour problems, narrative coherence, and expression of prosocial themes up to two years later. These findings make a small contribution to the case that investing in the earliest years can reap long-term rewards. Further investigations of the mechanisms underlying these effects may enhance our understanding of the ways in which parenting contributes not only to children's observable behaviours but also aspects of their internal functioning. Longer-term follow-up of these children would also allow us to consider if these positive effects extend to middle childhood and beyond.

The findings also have implications for our understanding of the use of the VIPP-SD programme. To my knowledge, this is the first study of VIPP-SD that suggests a positive effect on a measure of children's internal functioning. However, overall, the effects appeared to be of a small magnitude. The effects on parental sensitivity in particular (*d*s .11 and .23) were smaller than in previous studies of the programme's use with children recruited based on behavioural difficulties (children with externalising difficulties: $d = .34$; Juffer, Bakermans-Kranenburg, et al., 2017b). This may represent a genuine voltage drop in the effects of the programme when delivered through public health services. Alternatively, methodological issues related to the coding of parental sensitivity may have reduced the reliability of findings. The current findings would be enhanced by an extension of the assessment of changes to sensitivity to the remainder of the HSHS sample, as well as a consideration of findings through the lens of differential susceptibility. Finally, an examination of the long-term cost-effectiveness of delivering the intervention through public health settings would enable us to

explore the investment required for what are relatively small (but still potentially powerful) behavioural changes. Costs of enduring behaviour problems for education, health, and social care services are likely to accrue as children get older. A robust economic evaluation of the cost-related impacts of intervention-driven behavioural changes over long-term trajectories could make a powerful contribution to the case for early intervention.

6.5.3 A single intervention will not address the dynamic needs of families

This thesis provides supportive evidence for the potential of early intervention in improving child outcomes. However, the findings were mixed and where there were positive intervention effects they were of a small magnitude. Previous analysis of the full HSHS sample also revealed that the magnitude of the programme's effect on children's behaviour problems diminished by about 15% over two years, with a decrease in effect size of $d = 0.03$ between the five- and 24-month follow-up assessment (O'Farrelly, Barker, et al., 2021). This by no means indicates that such interventions are not worth our investment. But what it does suggest is that addressing the complex and varied needs of families requires more than just a single programme delivered in isolation during toddlerhood. The discovery of an all-powerful, universal remedy or 'magic bullet' is unlikely. Instead, we must focus on designing and testing multiple feasible, flexible, and scalable prevention and intervention strategies. This necessitates consistent and reliable investment for the staffing, training, funding, and organisational capacity required for such development and delivery. Combined, these efforts could enable us to offer families a package of evidence-based supports that could be tailored to their needs to promote their functioning through the life course.

One relatively new and particularly innovative approach that may make such packages of support more realistic to deliver involves 'components' or 'common elements' analysis. This explores the characteristics of interventions in an effort to isolate the features, practices, and modalities that prove to be most consistently effective at promoting positive outcomes. The aim of this analytical approach is to 'pick out the golden threads' (Ghiara & Clarke, 2020). The idea is that identifying such components may allow for the repackaging of briefer, personalised, more targeted supports. Such approaches have been utilised in studies of adolescent mental health (Skeen et al., 2019), child mental health (Chorpita & Daleiden, 2009), and in early childhood education in the UK (Clarke, Baker, McBride, & Ramchandani, 2020). However, there is a gap in such an examination for programmes targeted at families with infants and toddlers. In the Chorpita and Daleiden (2009) review, 92% of studies involved children aged 4-19 years. Knowing what common elements of programmes work

best for families in their earliest years could allow for the personalised delivery of supports, even when funding and resource is low.

6.5.4 Systemic and entrenched issues obstruct families' capacity to access, engage in, and benefit from support

However, even with robustly tested, effective, and scalable interventions (or components of interventions) there are more systemic and entrenched barriers to families accessing, engaging with, and benefiting from supports that must be acknowledged. Epidemiological research indicates that socioeconomic inequalities in mental health and wellbeing are widening, and that experiencing such disparities in the early years has lifelong impacts (Marmot, Allen, Boyce, Goldblatt, & Morrison, 2020). In 2017-18, four million (30%) children in England were living in poverty (Corlett, Clarke, & Tomlinson, 2017). At the time, this figure was predicted to rise to 36% by 2021 (Hood & Waters, 2017) – a projection made long before we realised the appalling effects the COVID-19 pandemic would have on amplifying the inequalities between the most and least privileged in society (Marmot & Allen, 2020; Office for National Statistics, 2020a). High-quality supports delivered early are considered to be the most clinically effective, cost-effective, and equity-effective way to disrupt these disparities (Marmot, 2010). However, services that offer vital support to families in the five years before formal schooling begins have faced substantial cuts over the last decade (Marmot et al., 2020). The allocation of funding to early intervention services for children and young people is reported to have fallen by 46% between 2011 and 2018 (from £3.5bn to £1.9bn; Action for Children, National Children's Bureau, and the Children's Society, 2016). An estimated 30% of children's centres have closed, with local authorities (and consequently children and families) in the most deprived areas disproportionately affected by these cuts with funding falling twice as fast (Smith, Sylva, Smith, Sammons, & Omonigho, 2018). Recent reports have also revealed that early years education settings and childcare providers have been especially hard-hit by the financial instability brought about by the COVID-19 pandemic, in the context of spending on childcare provision for two-year-olds already dropping by 9% between 2018-19 and 2019-20 (Britton, Farquharson, Sibieta, Tahir, & Waltmann, 2020).

Social and economic hardship directly impacts on parents' capacity to parent in the way they would like to (Magnuson & Duncan, 2002; Morrison et al., 2014). While parenting programmes have the potential for positive change, reducing poverty has been proposed as a more far-reaching strategy for improving children's outcomes (Marmot et al., 2020). The

effects of poverty are only compounded by other inequalities. Poverty rates are higher for children living in lone-parent families (49%), for children in large families (3+ children; 47%), and for children from Black, Asian, and Minority Ethnic groups (46%) (Department for Work and Pensions, 2021). Those living in poverty are also more likely to have experienced adverse childhood experiences (Steele et al., 2016). Such experiences have been estimated to account for 30% of adult mental health issues (Kessler et al., 2010). Detecting and understanding the effects of such inequalities is far beyond the scope of this thesis. But the presence of unequal effects of socioeconomic hardship on participating families is undeniable, even if it went unmeasured in these studies. UK Government initiatives that have actively sought to reduce the effects of these disparities have shown marked impacts on primary school aged children attending London schools (e.g., Teach First, London Challenge; Baars et al., 2014). With sufficient and stable investment, nurturing and extending comparable initiatives for new and expectant parents, infants, toddlers, and pre-schoolers may allow us to see such positive effects at an even earlier life stage. Such proactive action is particularly salient at this time given that early reports have indicated that it is the most deprived children, and those from Black, Asian, and Minority Ethnic families, who are most likely to have been affected by the loss of learning and increased educational inequalities that have emerged through lockdowns since March 2020 (Britton et al., 2020).

The recent publication of the first phase of the Early Years Healthy Development Review chaired by Andrea Leadsom (Department of Health and Social Care, 2021) outlined a vision for supporting babies, children, and families through the 1,001 critical days from conception to age two. The review sets out a programme of work that envisions the roll-out of a joined-up, multidisciplinary 'Start for Life' system to ensure families have access to the services and support they need. However, as stated in the report, "words on the page won't change what happens in communities and won't improve the support families receive; it's what we do next that matters" (p.11). Ensuring this vision is implemented effectively, using evidence-based pathways of support, and close consultation with families and service providers themselves, through processes underpinned by significant and sustained investment, could go some way in addressing the needs of children and their families at the very outset of development. Key to this is the formation of an economic case for further investment in the earliest years, something that requires an appraisal of the impacts of existing spending. Encouragingly, this need forms the focus of the next phase of the ongoing Early Years Healthy Development Review, the outcomes of which may be key to obtaining funding to

begin to address the social and health inequalities facing families today. The findings of this thesis, the wider HSHS study, and other investigations of early intervention and child development have the potential to make small but valuable evidence-based contributions to these endeavours.

It is also critical to consider the current landscape of health service provision in the discussion of this research. The intervention studied in this thesis was largely delivered through health visiting services. A survey conducted at the end of 2019 (Institute of Health Visiting, 2020a) revealed that these services continue to be on the receiving end of significant cuts within a "dwindling and all too frequently demoralised workforce" (pg.2). In 2015, the responsibility for commissioning public health services for children aged zero to five was transferred from the NHS to local authorities as part of the Healthy Child Programme (Department of Health and Social Care, 2009). The Institute of Health Visiting (2020a) survey reported that between 2015 and 2019, there was a 24% decrease in the numbers of health visitors in England. In total, 43% of health visitors were found to each be responsible for 400 – 1000+ children. In the last two years, health visitors reported a 78% perceived increase in demand for child behaviour problems and 83% for perinatal mental health problems, but 38% rated the care they were able to provide as 'inadequate' or 'poor', and 48% reported feeling so overstretched that they feared a tragedy at some point. A follow-up survey conducted in November 2020, found that in the first waves of the COVID-19 pandemic, in some areas 50% of health visitors were redeployed, and only 17% and 10% of health visitors were able to offer families their one- and two-year checks respectively (Institute of Health Visiting, 2020b). The findings of this thesis, coupled with earlier analysis of the Healthy Start, Happy Start study (O'Farrelly, Barker, et al., 2021; O'Farrelly, Watt, et al., 2021), provide important evidence of the power of health visitors delivering early intervention for young children and families. To best serve families though, these services need adequate and sustained investment, funding, and capacity to deliver high-quality and evidence-based supports to caregivers and children as they move through the family life course. In the current context, an uncertainty remains around whether services would have the capacity and practitioner-power to deliver a programme like VIPP-SD as part of routine practice.

6.6 Conclusion

This thesis has highlighted the feasibility and enhanced value that can come from including children themselves in research about their lives. The use of a play-based narrative

story stem tool proved to be engaging to almost all children and provided data that corresponded with parent- and teacher-reports but also offered us unique insight into children's development. Not only this, it allowed us to detect effects of an intervention on aspects of children's narrative representations, and the important predictive role of early parental sensitivity and children's early behaviour problems on these outcomes. The findings make a unique contribution to the potential of early parenting interventions to not only improve children's behaviour but also elements of their internal representations, as measured through the story stem tool. However, the accessibility, feasibility, and scalability of such interventions must be considered in the wider contexts in which children, families, and services find themselves.

References

Abidin, R., Flens, J. R., & Austin, W. G. (2006). The parenting stress index. In R. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 297-328). New York: Routledge.

Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science, 15*(2), 94-98.

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232.

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families.

Action for Children National Children's Bureau and the Children's Society. (2016). Losing in the long run. Trends in early intervention funding. Action for Children. Retrieved from https://www.actionforchildren.org.uk/media/5826/losing_in_the_long_run.pdf

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. F. (1971). Individual differences in strange-situational behaviour of one-year-olds. In H. R. Schaffer (Ed.), *The origins of human social relations* (pp. 17-58). London: Academic Press.

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. F. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In M. P. M. Richards (Ed.), *The integration of a child into a social world* (pp. 99-135). London: Cambridge University Press.

Ainsworth, M. D. S., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Albanese, A. M., Russo, G. R., & Geller, P. A. (2019). The role of parental self-efficacy in parent and child well-being: A systematic review of associated outcomes. *Child: Care, Health and Development, 45*(3), 333-363.

Alink, L. R., Mesman, J., van Zeijl, J., Stolk, M. N., Juffer, F., Koot, H. M., . . . van IJzendoorn, M. (2006). The early childhood aggression curve: Development of physical aggression in 10-to 50-month-old children. *Child Development, 77*(4), 954-966.

Allen, G. (2011). *Early intervention: The next steps*. London: The Stationery Office.

Allen, K., Marlow, R., Edwards, V., Parker, C., Rodgers, L., Ukoumunne, O. C., . . . Ford, T. (2018). 'How I Feel About My School': The construction and validation of a measure of wellbeing at school for primary school children. *Clinical Child Psychology and Psychiatry, 23*(1), 25-41.

American Psychiatric Association. (2004). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Anderson, K. E., Lytton, H., & Romney, D. M. (1986). Mothers' interactions with normal and conduct-disordered boys: Who affects whom? *Developmental Psychology, 22*(5), 604-609.

Angold, A., & Costello, E. J. (1993). Depressive comorbidity in children and adolescents: empirical, theoretical, and methodological issues. *American Journal of Psychiatry, 150*(12), 1779-1791.

Antal, H., Hossain, M. J., Hassink, S., Henry, S., Fuzzell, L., Taylor, A., & Wysocki, T. (2015). Audio-video recording of health care encounters for pediatric chronic conditions: observational reactivity and its correlates. *Journal of Pediatric Psychology, 40*(1), 144-153.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*(7), 602-614.

Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5*(2), 137-144.

Arseneault, L., Kim-Cohen, J., Taylor, A., Caspi, A., & Moffitt, T. E. (2005). Psychometric evaluation of 5-and 7-year-old children's self-reports of conduct problems. *Journal of Abnormal Child Psychology, 33*(5), 537-550.

Aspland, H., & Gardner, F. (2003). Observational measures of parent-child interaction: An introductory review. *Child and Adolescent Mental Health, 8*(3), 136-143.

Atkinson, L., Niccols, A., Paglia, A., Coolbear, J., Parker, K. C. H., Poulton, L., . . . Sitarenios, G. (2000). A meta-analysis of time between maternal sensitivity and attachment assessments: Implications for internal working models in infancy/toddlerhood. *Journal of Social and Personal Relationships, 17*(6), 791-810.

Baars, S., Barnardes, E., A, E., Malortie, A., McAleavy, T., McInerney, L., . . . Riggall, A. (2014). *Lessons from London Schools: Investigating the success*. Reading: CfBT Education Trust.

Bachmann, C. J., Beecham, J., O'Connor, T. G., Briskman, J., & Scott, S. (2021). A good investment: Longer-term cost savings of sensitive parenting in childhood. *Journal of Child Psychology and Psychiatry*. doi:10.1111/jcpp.13461

Bakermans-Kranenburg, M., van IJzendoorn, M., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*(2), 195-215.

Bakermans-Kranenburg, M., van IJzendoorn, M., Pijlman, F. T., Mesman, J., & Juffer, F. (2008). Experimental evidence for differential susceptibility: Dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Developmental Psychology, 44*(1), 293-300.

Bakermans-Kranenburg, M., Lotz, A., Alyousefi-van Dijk, K., & van IJzendoorn, M. (2019). Birth of a father: Fathering in the first 1,000 days. *Child Development Perspectives, 13*(4), 247-253.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A., & Walters, R. (1963). *Social learning and personality development*. New York: Holt, Rinehart, & Winston.

Barker, B., Iles, J. E., & Ramchandani, P. G. (2017). Fathers, fathering and child psychopathology. *Current Opinion in Psychology, 15*, 87-92.

Barkley, R. A. (1988). The effects of methylphenidate on the interactions of preschool ADHD children with their mothers. *Journal of the American Academy of Child and Adolescent Psychiatry, 27*(3), 336-341.

Barlow, J., Bennett, C., Midgley, N., Larkin, S. K., & Wei, Y. (2015). Parent-infant psychotherapy for improving parental and infant mental health. *Cochrane Database of Systematic Reviews*(1). doi:10.1002/14651858.CD010534.pub2

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 6*(51), 1173-1182.

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development, 37*(4), 887-907.

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genetic Psychology Monographs*.

Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence, 11*(1), 56-95.

Baumrind, D. (2012). Differentiating between confrontive and coercive kinds of parental power-assertive disciplinary practices. *Human Development, 55*(2), 35-51.

Beijersbergen, M. D., Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2012). Remaining or becoming secure: parental sensitive support predicts attachment continuity from infancy to adolescence in a longitudinal adoption study. *Developmental Psychology, 48*(5), 1277-1282.

Belden, A. C., Sullivan, J. P., & Luby, J. L. (2007). Depressed and healthy preschoolers' internal representations of their mothers' caregiving: Associations with observed caregiving behaviors one year later. *Attachment & human development, 9*(3), 239-254.

Belsky, J., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science, 16*(6), 300-304.

Belsky, J., & Fearon, R. M. P. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: does continuity in development depend upon continuity of caregiving? *Attachment & human development, 4*(3), 361-387.

Belsky, J., & Fearon, R. M. P. (2008). Precursors of attachment security. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 295-316). New York, NY: Guilford Press.

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin, 135*(6), 885-908.

Belsky, J., & van IJzendoorn, M. (2017). Genetic differential susceptibility to the effects of parenting. *Current Opinion in Psychology, 15*, 125-130.

Beltz, A., Blakemore, J., & Berenbaum, S. A. (2013). Sex differences in brain and behavioral development. In P. Rakic & J. Rubenstein (Eds.), *Comprehensive developmental neuroscience* (Vol. 3, pp. 467-499). Oxford: Elsevier.

Bennetts, S. K., Mensah, F. K., Green, J., Hackworth, N. J., Westrupp, E. M., & Reilly, S. (2017). Mothers' experiences of parent-reported and video-recorded observational assessments. *Journal of Child and Family Studies, 26*(12), 3312-3326.

Beresford, C., Robinson, J. L., Holmberg, J., & Ross, R. G. (2007). Story stem responses of preschoolers with mood disturbances. *Attachment & human development, 9*(3), 255-270.

Bergen, P. V., Salmon, K., Dadds, M. R., & Allen, J. (2009). The effects of mother training in emotion-rich, elaborative reminiscing on children's shared recall and emotion knowledge. *Journal of Cognition and Development, 10*(3), 162-187.

Bernard, K., Dozier, M., Bick, J., Lewis-Morrarty, E., Lindhiem, O., & Carlson, E. (2012). Enhancing attachment organization among maltreated children: Results of a randomized clinical trial. *Child Development, 83*(2), 623-636.

Berzenski, S. R., & Yates, T. M. (2017). The differential influences of parenting and child narrative coherence on the development of emotion recognition. *Developmental Psychology, 53*(10), 1912.

Bettmann, J. E., & Lundahl, B. W. (2007). Tell me a story: A review of narrative assessments for preschoolers. *Child and Adolescent Social Work Journal, 24*(5), 455-475.

Biedzio, D., & Wakschlag, L. S. (2019). Developmental emergence of disruptive behaviors beginning in infancy: Delineating normal:abnormal boundaries to enhance early identification. In C. H. Zeanah (Ed.), *Handbook of infant mental health* (4th ed., pp. 407-425). New York: Guilford Press.

Biringen, Z., & Easterbrooks, M. A. (2012). Emotional availability: concept, research, and window on developmental psychopathology. *Development and Psychopathology, 24*(1), 1-8.

Biringen, Z., Robinson, J. L., & Emde, R. N. (2000). The emotional availability scales: An abridged infancy/early childhood version. *Attachment & human development, 2*(2), 256-270.

Birmingham, R. S., Bub, K., & Vaughn, B. (2017). Parenting in infancy and self-regulation in preschool: An investigation of the role of attachment history. *Attachment & human development, 19*(2), 107-129.

Blaisdell, C., Arnott, L., Wall, K., & Robinson, C. (2019). Look who's talking: using creative, playful arts-based methods in research with young children. *Journal of Early Childhood Research, 17*(1), 14-31.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: John Wiley & Sons.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science and Medicine, 75*(12), 2299-2306.

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language, 24*(3), 267-304.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Directions for Child and Adolescent Development, 1989*(43), 49-61.

Bovenschen, I., Lang, K., Zimmermann, J., Förthner, J., Nowacki, K., Roland, I., & Spangler, G. (2016). Foster children's attachment behavior and representation: Influence of children's pre-placement experiences and foster caregiver's sensitivity. *Child Abuse and Neglect, 51*, 323-335.

Bowlby, J. (1969). Attachment and loss v. 3 (Vol. 1). In: Random House. Furman, W., & Buhrmester, D.(2009). Methods and measures: The ….

Bowlby, J. (1969/1982). *Attachment and loss: Vol. 1. Attachment.* (2nd ed.). New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation: Anxiety and anger.* New York: Basic Books.

Brace, N., Kemp, R., & Snelgar, R. (2012). *SPSS for Psychologists* (5th ed.). New York, NY: Palgrave MacMillan.

Branger, M. C., Emmen, R. A., Woudstra, M.-l. J., Alink, L. R., & Mesman, J. (2019). Context matters: Maternal and paternal sensitivity to infants in four settings. *Journal of Family Psychology, 33*(7), 851-856.

Braungart-Rieker, J. M., Zentall, S., Lickenbrock, D. M., Ekas, N. V., Oshio, T., & Planalp, E. (2014). Attachment in the making: Mother and father sensitivity and infants' responses during the Still-Face Paradigm. *Journal of Experimental Child Psychology, 125*, 63-84.

Bretherton, I. (1985). Attachment theory: Retrospect and prospect. *Monographs of the Society for Research in Child Development, 50*(1-2), 3-35.

Bretherton, I. (1987). New perspectives on attachment relations: Security, communication and internal working models. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 1061-1100). New York: Wiley.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18*(6), 906-921.

Bretherton, I., Biringen, Z., Ridgeway, D., Maslin, C., & Sherman, M. (1989). Attachment: The parental perspective. *Infant Mental Health Journal, 10*(3), 203-221.

Bretherton, I., & Munholland, K. A. (2008). Internal working models in attachment relationships: Elaborating a central construct in attachment theory. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 102-127). New York: The Guilford Press.

Bretherton, I., & Oppenheim, D. (2003). The MacArthur Story Stem Battery: Development, administration, reliability, validity, and reflections about meaning. In R. Emde, D. P. Wolf, & D. Oppenheim (Eds.), *Revealing the inner worlds of young children: The MacArthur Story Stem Battery and parent-child narratives* (pp. 55-80). Oxford: Oxford University Press.

Bretherton, I., Oppenheim, D., Buchsbaum, H., & Emde, R. (1990). *The MacArthur Story Stem Battery (MSSB)*. University of Wisconsin-Madison, Madison: Unpublished manuscript.

Bretherton, I., Prentiss, C., & Ridgeway, D. (1990). Family relationships as represented in a story-completion task at thirty-seven and fifty-four months of age. *New Directions for Child and Adolescent Development, 1990*(48), 85-105.

Bretherton, I., Ridgeway, D., & Cassidy, J. (1990). Assessing internal working models of the attachment relationship: An attachment story completion task for 3-year-olds. In M. Greenberg, D. Cicchetti, & M. Cummings (Eds.), *Attachment in the preschool years: Theory, research, and intervention*. Chicago, IL: The University of Chicago Press.

Britto, P. R., Lye, S. J., Proulx, K., Yousafzai, A. K., Matthews, S. G., Vaivada, T., . . . Fernald, L. C. (2017). Nurturing care: promoting early childhood development. *Lancet, 389*(10064), 91-102.

Britton, J., Farquharson, C., Sibieta, L., Tahir, I., & Waltmann, B. (2020). *2020 annual report on education spending in England*. London: Institute for Fiscal Studies.

Brody, L. R. (1993). On understanding gender differences in the expression of emotion: Gender roles, socialization, and language. In S. L. Ablon, D. P. Brown, E. J. Khantzian, & J. E. Mack (Eds.), *Human feelings: Explorations in affect development and meaning* (pp. 87-121). Hillsdale, NJ: The Analytic Press.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U., & Morris, A. (1998). The ecology of developmental processes. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 993-1028). Hoboken, NJ: John Wiley.

Bronfman, E., Madigan, S., & Lyons-Ruth, K. (2009). *The atypical behavior instrument for assessment and classification*. Cambridge, MA: Harvard Medical School.

Buchsbaum, H. K., & Emde, R. N. (1990). Play narratives in 36-month-old children: Early moral development and family relationships. *The Psychoanalytic Study of the Child, 45*(1), 129-155.

Buchsbaum, H. K., Toth, S. L., Clyman, R. B., Cicchetti, D., & Emde, R. N. (1992). The use of a narrative story stem technique with maltreated children: Implications for theory and practice. *Development and Psychopathology, 4*(4), 603-625.

Byford, S., Cary, M., Barrett, B., Aldred, C. R., Charman, T., Howlin, P., . . . McConachie, H. (2015). Cost-effectiveness analysis of a communication-focused therapy for pre-school children with autism: Results from a randomised controlled trial. *BMC Psychiatry, 15*(1), 316-328.

Cabrera, N. J., Volling, B. L., & Barr, R. (2018). Fathers are parents, too! Widening the lens on parenting for children's development. *Child Development Perspectives, 12*(3), 152-157.

Capaldi, D. (1991). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: I. Familial factors and general adjustment at Grade 6. *Development and Psychopathology, 3*(3), 277-300.

Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology and Psychiatry, 45*(1), 109-134.

Cassibba, R., Castoro, G., Costantino, E., Sette, G., & van IJzendoorn, M. (2015). Enhancing maternal sensitivity and infant attachment security with video feedback: An exploratory study in Italy. *Infant Mental Health Journal, 36*(1), 53-61.

Cassidy, J., & Marvin, R. S. (1992). *Attachment organization in preschool children: Procedures and coding manual*. University of Virginia: Unpublished manuscript

Census. (2011a). Detailed ethnicity by age and sex ward tools. Retrieved from https://data.london.gov.uk/dataset/detailed-ethnicity-by-age---sex-ward-tools---2011-census--

Census. (2011b). Population estimates by sex, age and highest level of qualification. Retrieved from https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/populationestimatesbysexageandhighestlevelofqualification

Cents, R. A., Kok, R., Tiemeier, H., Lucassen, N., Székely, E., Bakermans-Kranenburg, M., . . . Verhulst, F. C. (2014). Variations in maternal 5-HTTLPR affect observed sensitive parenting. *Journal of Child Psychology and Psychiatry, 55*(9), 1025-1032.

Chang, L., Schwartz, D., Dodge, K. A., & McBride-Chang, C. (2003). Harsh parenting in relation to child emotion regulation and aggression. *Journal of Family Psychology, 17*(4), 598-606.

Children Act. (1989). c. 41. United Kingdom. Retrieved from https://www.legislation.gov.uk/ukpga/1989/41/contents

Choe, D. E., Olson, S. L., & Sameroff, A. J. (2013). The interplay of externalizing problems and physical and inductive discipline during childhood. *Developmental Psychology, 49*(11), 2029-2039.

Chorpita, B. F., & Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: Application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology, 77*(3), 566-579.

Chowdry, H., & Fitzsimons, P. (2016). *The cost of late intervention: EIF analysis 2016*. London: Early Intervention Foundation.

Christensen, P., & Prout, A. (2002). Working with ethical symmetry in social research with children. *Childhood, 9*(4), 477-497.

Clark, H., Coll-Seck, A. M., Banerjee, A., Peterson, S., Dalglish, S. L., Ameratunga, S., . . . Borrazzo, J. (2020). A future for the world's children? A WHO–UNICEF–Lancet Commission. *Lancet, 395*(10224), 605-658.

Clarke, A., Baker, S., McBride, T., & Ramchandani, P. (2020). The 'common elements' approach: Improving outcomes in early childhood education. Retrieved from https://www.nuffieldfoundation.org/project/common-elements-approach-improving-outcomes-early-childhood-education

Clarke, A. M., & Clarke, A. D. B. (2000). *Early experience and the life path*. London: Jessica Kingsley Publishers.

Clayman, M. L., & Wissow, L. S. (2004). Pediatric residents' response to ambiguous words about child discipline and behavior. *Patient Education and Counseling, 55*(1), 16-21. doi:10.1016/S0738-3991(03)00245-3

Cohen, J. (1969). *Statistical power analysis for the behavioural sciences*. New York, NY: Academic Press.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.
Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior, 24*(4), 385-396.
Coldwell, J., Pike, A., & Dunn, J. (2006). Household chaos – links with parenting and child behaviour. *Journal of Child Psychology and Psychiatry, 47*(11), 1116-1122.
Collishaw, S., Goodman, R., Ford, T., Rabe-Hesketh, S., & Pickles, A. (2009). How far are associations between child, family and community factors and child psychopathology informant-specific and informant-general? *Journal of Child Psychology and Psychiatry, 50*(5), 571-580.
Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York: Chapman & Hall.
Cooke, J. E., Eirich, R., Racine, N., Lyons-Ruth, K., & Madigan, S. (2020). Validation of the AMBIANCE-brief: An observational screening instrument for disrupted caregiving. *Infant Mental Health Journal, 41*(3), 299-312.
Corlett, A., Clarke, S., & Tomlinson, D. (2017). *The living standards audit 2017*. London: Resolution Foundation.
Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.
Costello, E. J., & Maughan, B. (2015). Annual research review: Optimal outcomes of child and adolescent mental illness. *Journal of Child Psychology and Psychiatry, 56*(3), 324-341.
Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.
Crittenden, P. (2001). *CARE-index manual*. Miami, FL: Family Relations Institute.
Crittenden, P. (2005). Using the CARE-Index for screening, intervention, and research. *Bindungsorientierte Ansätze in Der Praxis Der Frühförderung, 3*(24), 99-106. Retrieved from http://www.patcrittenden.com/images/CARE-Index.pdf
Crittenden, P. (2010). *CARE-Index. Infants. Coding manual*. Miami, FL: Unpublished manuscript.
Croft, S., Stride, C., Maughan, B., & Rowe, R. (2015). Validity of the Strengths and Difficulties Questionnaire in preschool-aged children. *Pediatrics, 135*(5), e1210-e1219.
Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S. M. (2017). Psychometric properties and normative data for the preschool Strengths and Difficulties Questionnaire in two-year-old children. *Journal of Abnormal Child Psychology, 45*(2), 345-357.

Dahlberg, A., Ghaderi, A., Sarkadi, A., & Salari, R. (2019). SDQ in the hands of fathers and preschool teachers - Psychometric properties in a non-clinical sample of 3–5-year-olds. *Child Psychiatry and Human Development, 50*(1), 132-141.

Dalgaard, N. T., Todd, B. K., Daniel, S. I., & Montgomery, E. (2016). The transmission of trauma in refugee families: Associations between intra-family trauma communication style, children's attachment security and psychosocial adjustment. *Attachment & human development, 18*(1), 69-89.

Danforth, J. S., Anderson, L., Barkley, R. A., & Stokes, T. F. (1991). Observations of parent-child interactions with hyperactive children: Research and clinical implications. *Clinical Psychology Review, 11*(6), 703-727.

Darbyshire, P., MacDougall, C., & Schiller, W. (2005). Multiple methods in qualitative research with children: More insight or just more? *Qualitative Research, 5*(4), 417-436.

Davies, P. T., Woitach, M. J., Winter, M. A., & Cummings, E. M. (2008). Children's insecure representations of the interparental relationship and their school adjustment: The mediating role of attention difficulties. *Child Development, 79*(5), 1570-1582.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin, 131*(4), 483-509.

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. (2013). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology, 9*, 123-149.

de Nijs, P. F. A., Ferdinand, R. F., de Bruin, E. I., Dekker, M. C. J., van Duijn, C. M., & Verhulst, D. C. (2004). Attention-deficit/hyperactivity disorder (ADHD): Parents' judgment about school, teachers' judgment about home. *European Child and Adolescent Psychiatry, 13*(5), 315-320.

De Wolff, M. S., & van IJzendoorn, M. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 68*(4), 571-591.

Dealy, J., Mudrick, H., & Robinson, J. (2019). Children's narrative story stem responses: Contributions of executive functioning and language proficiency to relationship representations. *Social Development, 28*(1), 168-185.

Deighton, J., Croudace, T., Fonagy, P., Brown, J., Patalay, P., & Wolpert, M. (2014). Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: A review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health, 8*(1), 1-14.

DeKlyen, M., & Greenberg, M. (2016). Attachment and psychopathology in childhood. In J. S. Cassidy, PR (Ed.), *Handbook of attachment theory, research, and clinical applications* (3rd ed., pp. 637-665). New York: Guilford Press.

Denham, S. A., Warren-Khot, H. K., Bassett, H. H., Wyatt, T., & Perna, A. (2012). Factor structure of self-regulation in preschoolers: Testing models of a field-based assessment for predicting early school readiness. *Journal of Experimental Child Psychology, 111*(3), 386-404.

Dent, V., & Goodman, G. (2020). Representations of attachment security, attachment avoidance, and gender in Ugandan children. *Attachment & human development*, 1-30. doi:10.1080/14616734.2020.1830480

Department for Work and Pensions. (2021). Households below average income: for financial years ending 1995 to 2020. Retrieved from https://www.gov.uk/government/statistics/households-below-average-income-for-financial-years-ending-1995-to-2020

Department of Health and Social Care. (2009). Healthy Child Programme: Pregnancy and the first 5 years of life. Retrieved from https://www.gov.uk/government/publications/healthy-child-programme-pregnancy-and-the-first-5-years-of-life

Department of Health and Social Care. (2021). *The best start for life: a vision for the 1,001 critical days*. London: HM Government.

Department of Health and Social Care and Department for Education. (2017). *Transforming children and young people's mental health provision: A green paper*. London: Department of Health and Department for Education.

DePasquale, C. E., & Gunnar, M. R. (2020). Parental Sensitivity and Nurturance. In J. Brooks-Gunn & A. Aizer (Eds.), *Future of children: Supporting development from three trimesters to three years* (pp. 53-70). Princeton, NJ: Princeton-Brookings.

Dockett, S., Einarsdottir, J., & Perry, B. (2009). Researching with children: Ethical tensions. *Journal of Early Childhood Research, 7*(3), 283-298.

Dockett, S., & Perry, B. (2005). Children's drawings: Experiences and expectations of school. *International Journal of Equity and Innovation in Early Childhood, 3*(2), 77-89.

Doi, Y., Ishihara, K., & Uchiyama, M. (2014). Reliability of the strengths and difficulties questionnaire in Japanese preschool children aged 4–6 years. *Journal of Epidemiology, 24*(6), 514-518.

Donelan-McCall, N. (2017). Parenting and home visiting interventions. In E. Votruba-Drzal & E. Dearing (Eds.), *The Wiley handbook of early childhood development programs, practices, and policies* (pp. 310-329). Oxford: Wiley Blackwell.

Doyle, O., Harmon, C. P., Heckman, J. J., & Tremblay, R. E. (2009). Investing in early human development: Timing and economic efficiency. *Economics and Human Biology, 7*(1), 1-6.

Dubois-Comtois, K., Cyr, C., & Moss, E. (2011). Attachment behavior and mother-child conversations as predictors of attachment representations in middle childhood: A longitudinal study. *Attachment & human development, 13*(4), 335-357.

Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice, 7*(4), 435-453.

Duncombe, M. E., Havighurst, S. S., Holland, K. A., & Frankling, E. J. (2012). The contribution of parenting practices and parent emotion factors in children at risk for disruptive behavior disorders. *Child Psychiatry and Human Development, 43*(5), 715-733.

Dunifon, R. E., Near, C. E., & Ziol-Guest, K. M. (2018). Backup parents, playmates, friends: Grandparents' time with grandchildren. *Journal of Marriage and Family, 80*(3), 752-767.

Dunn, J., & Hughes, C. (2001). "I got some swords and you're dead!": Violent fantasy, antisocial behavior, friendship, and moral sensibility in young children. *Child Development, 72*(2), 491-505.

Duschinsky, R., Bakkum, L., Mannes, J. M. M., Skinner, G. C. M., Turner, M., Mann, A., . . . Beckwith, H. (2021). Six attachment discourses: Convergence, divergence and relay. *Attachment & human development*, 1-20. doi:10.1080/14616734.2021.1918448

Dykas, M. J., & Cassidy, J. (2011). Attachment and the processing of social information across the life span: Theory and evidence. *Psychological Bulletin, 137*(1), 19-46.

Egeland, B., Erickson, M. F., Clemenhagen-Moon, J., Hiester, M. K., & Korfmacher, J. (1990). *24 month tools coding manual: Project STEEP: Revised 1990 from mother–child project scales*. Department of Psychology, University of Minnesota: Unpublished manuscript.

Egger, H. L., & Angold, A. (2004). The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In R. Del Carmen-Wiggins & A. Carter (Eds.), *Handbook of infant, toddler, and preschool mental health assessment* (pp. 223-243). New York: Oxford University Press.

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, 47*(3-4), 313-337.

Einarsdóttir, J. (2007). Research with children: methodological and ethical challenges. *European Early Childhood Education Research Journal, 15*(2), 197-211.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 646-718). New York: John Wiley & Sons.

Eisenberg, N., Losoya, S., Fabes, R. A., Guthrie, I. K., Reiser, M., Murphy, B., . . . Padgett, S. J. (2001). Parental socialization of children's dysregulated expression of emotion and externalizing problems. *Journal of Family Psychology, 15*(2), 183-205.

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology, 23*(1), 7-28.

Emde, R. N., Wolf, D. P., & Oppenheim, D. (2003). *Revealing the inner worlds of young children: The MacArthur Story Stem Battery and parent-child narratives*. Oxford: Oxford University Press.

Emond, R. (2002). *Learning from their lessons: A study of young people in residential care and their experiences of education*. Dublin, Ireland: The Children's Research Centre.

Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. *Monographs of the Society for Research in Child Development, 50*(1-2), 147-166.

Eshel, N., Daelmans, B., Mello, M. C. d., & Martines, J. (2006). Responsive parenting: Interventions and outcomes. *Bulletin of the World Health Organization, 84*, 992-999.

European Commission. (2021). EU strategy on the right of the child (2021-24). Retrieved from https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52021DC0142

Euser, S., Vrijhof, C. I., Van den Bulk, B. G., Vermeulen, R., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2021). Video-feedback promotes sensitive limit-setting in parents of twin preschoolers: a randomized controlled trial. *BMC Psychology, 9*(1), 1-14.

Ezpeleta, L., Granero, R., de la Osa, N., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the Strengths and Difficulties Questionnaire in 3-year-old preschoolers. *Comprehensive Psychiatry, 54*(3), 282-291.

Fagan, J., Day, R., Lamb, M. E., & Cabrera, N. J. (2014). Should researchers conceptualize differently the dimensions of parenting for fathers and mothers? *Journal of Family Theory & Review, 6*(4), 390-405.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Fearon, R. M. P., Bakermans-Kranenburg, M., van IJzendoorn, M., Lapsley, A. M., & Roisman, G. I. (2010). The significance of insecure attachment and disorganization in the development of children's externalizing behavior: A meta-analytic study. *Child Development, 81*(2), 435-456.

Fearon, R. M. P., & Belsky, J. (2011). Infant–mother attachment and the growth of externalizing problems across the primary-school years. *Journal of Child Psychology and Psychiatry, 52*(7), 782-791.

Feeney, B. C., & Woodhouse, S. S. (2016). Caregiving. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 827-852). New York; London: Guilford Press.

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development, 71*(2), 310-322.

Feldman, R. (1998). *Coding Interactive Behaviour (CIB) Manual*. Unpublished manuscript: Bar-Ilan University

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., . . . Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 174-185.

Fergusson, D. M., Horwood, L., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*(8), 837-849.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: SAGE publications.

Fink, E., Mareva, S., & Gibson, J. L. (2020). Dispositional playfulness in young children: A cross-sectional and longitudinal examination of the psychometric properties of a new child self-reported playfulness scale and associations with social behaviour. *Infant and Child Development, 29*(4), e2181.

Fiorentino, L., & Howe, N. (2004). Language competence, narrative ability, and school readiness in low-income preschool children. *Canadian Journal of Behavioural Science, 36*(4), 280-294.

Flewitt, R. (2005). Conducting research with young children: Some ethical considerations. *Early Child Development and Care, 175*(6), 553-565.

Foley, S., Álvarez, C., McCarthy, J., & Hughes, C. (2020). Two's company, three's a crowd? Maternal and paternal talk about their infant differs in associations with wellbeing, couple relationship quality, and caregiving sensitivity. *Frontiers in Psychiatry, 11*, 578632. doi:10.3389/fpsyt.2020.578632

Fonagy, P., Gergely, G., & Target, M. (2007). The parent–infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry, 48*(3-4), 288-328.

Fonagy, P., Sleed, M., & Baradon, T. (2016). Randomized controlled trial of parent–infant psychotherapy for parents with mental health problems and young infants. *Infant Mental Health Journal, 37*(2), 97-114.

Fonagy, P., Steele, H., & Steele, M. (1991). Maternal representations of attachment during pregnancy predict the organization of infant-mother attachment at one year of age. *Child Development, 62*(5), 891-905.

Ford, K., Sankey, J., & Crisp, J. (2007). Development of children's assent documents using a child-centred approach. *Journal of Child Health Care, 11*(1), 19-28.

Fraiberg, S. (1980). *Clinical studies in infant mental health: The first year of life*. London: Tavistock.

Fraley, R. C., Roisman, G. I., & Haltigan, J. D. (2013). The legacy of early experiences in development: formalizing alternative models of how early experiences are carried forward over time. *Developmental Psychology, 49*(1), 109-126.

Friedli, L., & Parsonage, M. (2007). Building an economic case for mental health promotion: Part I. *Journal of Public Mental Health, 6*(3), 14-23.

Futh, A., O'Connor, T. G., Matias, C., Green, J., & Scott, S. (2008). Attachment narratives and behavioral and emotional symptoms in an ethnically diverse, at-risk sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(6), 709-718.

Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science, 3*(2), 206-215.

Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review, 3*(3), 185-198.

Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child and Adolescent Psychology, 39*(4), 568-580.

Gardner, F., Leijten, P., Melendez-Torres, G., Landau, S., Harris, V., Mann, J., . . . Scott, S. (2019). The earlier the better? Individual participant data and traditional meta-analysis of age effects of parenting interventions. *Child Development, 90*(1), 7-19.

Gardner, F., & Shaw, D. S. (2008). Behavioral problems of infancy and preschool children (0–5). In M. Rutter, D. Bishop, D. S. Pine, S. Scott, J. Stevenson, A. Taylor, & A.

Thapar (Eds.), *Rutter's child and adolescent psychiatry* (5th ed., pp. 882-892). Oxford: Blackwell Publishing.

George, C., Kaplan, N., & Main, M. (1996). *The Adult Attachment Interview*. University of California, Berkeley: Unpublished manuscript.

George, D., & Mallery, M. (2019). *SPSS for Windows step by step: A simple guide and reference* (6th ed.). New York: Routledge.

Gershoff, E. T., & Grogan-Kaylor, A. (2016). Spanking and child outcomes: Old controversies and new meta-analyses. *Journal of Family Psychology, 30*(4), 453-469.

Getz, J. A., Goldman, J. A., & Corsini, D. A. (1984). Interpersonal problem solving in preschool children: A comparison of assessment procedures using two-dimensional versus three-dimensional stimuli. *Journal of Applied Developmental Psychology, 5*(4), 293-304.

Ghiara, V., & Clarke, A. (2020). Picking out the golden threads: A common elements approach to supporting children's development. Retrieved from https://www.eif.org.uk/blog/picking-out-the-golden-threads-a-common-elements-approach-to-supporting-childrens-development

Giannopoulou, I. (2003). *Preschool hyperactivity: Development of observational and self-regulation measures for detecting hyperkinetic disorder in preschool children.* (PhD thesis). King's College London, London, UK. Retrieved from https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.404417

Gloger-Tippelt, G., Gomille, B., Koenig, L., & Vetter, J. (2002). Attachment representations in 6-year-olds: Related longitudinally to the quality of attachment in infancy and mothers' attachment representations. *Attachment & human development, 4*(3), 318-339.

Goldsmith, H., & Alansky, J. A. (1987). Maternal and infant temperamental predictors of attachment: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 55*(6), 805-816.

Goldstein, J. (2012). Play in children's development, health and well-being. Retrieved from https://www.ornes.nl/wp-content/uploads/2010/08/Play-in-children-s-development-health-and-well-being-feb-2012.pdf

Goldwyn, R., Stanley, C., Smith, V., & Green, J. (2000). The Manchester Child Attachment Story Task: Relationship with parental AAI, SAT and child behaviour. *Attachment & human development, 2*(1), 71-84.

Goodman, G., Aber, J. L., Berlin, L., & Brooks-Gunn, J. (1998). The relations between maternal behaviors and urban preschool children's internal working models of attachment security. *Infant Mental Health Journal, 19*(4), 378-393.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581–586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337-1345.

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Child Psychology, 27*(1), 17-24.

Green, H., McGinnity, Á., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in Great Britain, 2004*. Basingstoke, Hampshire: Palgrave Macmillan

Green, J., Charman, T., Pickles, A., Wan, M. W., Elsabbagh, M., Slonims, V., . . . Gliga, T. (2015). Parent-mediated intervention versus no intervention for infants at high risk of autism: A parallel, single-blind, randomised trial. *The Lancet Psychiatry, 2*(2), 133-140.

Green, J., Stanley, C., & Peters, S. (2007). Disorganized attachment representation and atypical parenting in young school age children with externalizing disorder. *Attachment & human development, 9*(3), 207-222.

Green, J., Stanley, C., Smith, V., & Goldwyn, R. (2000). A new method of evaluating attachment representations in young school-age children: The Manchester Child Attachment Story Task. *Attachment & human development, 2*(1), 48-70.

Greene, S., & Hill, M. (2005). Conceptual, methodological and ethical issues in researching children's experiences. In S. Greene & D. Hogan (Eds.), *Researching children's experience: Approaches and methods* (pp. 1-21). London: Sage.

Groeneveld, M. G., Vermeer, H. J., van IJzendoorn, M., & Linting, M. (2011). Enhancing home-based child care quality through video-feedback intervention: A randomized controlled trial. *Journal of Family Psychology, 25*(1), 86-96.

Groh, A. M., Fearon, R. M. P., Bakermans-Kranenburg, M., van IJzendoorn, M., Steele, R. D., & Roisman, G. I. (2014). The significance of attachment security for children's social competence with peers: A meta-analytic study. *Attachment & human development, 16*(2), 103-136.

Groh, A. M., Narayan, A. J., Bakermans-Kranenburg, M. J., Roisman, G. I., Vaughn, B. E., Fearon, R. P., & van IJzendoorn, M. H. (2017). Attachment and temperament in the early life course: A meta-analytic review. *Child Development, 88*(3), 770-795.

Groh, A. M., Roisman, G. I., van IJzendoorn, M., Bakermans-Kranenburg, M., & Fearon, R. M. P. (2012). The significance of insecure and disorganized attachment for children's internalizing symptoms: A meta-analytic study. *Child Development, 83*(2), 591-610.

Grossmann, K. E., & Grossmann, K. (2006). Attachment quality as an organizer of emotional and behavioral responses in a longitudinal perspective. In C. M. Parkes, J. Stevenson-Hinde, & P. Marris (Eds.), *Attachment across the life cycle* (pp. 101-122). New York, NY: Routledge.

Grossmann, K. E., Grossmann, K., & Waters, E. (2006). *Attachment from infancy to adulthood: The major longitudinal studies*. New York, NY: Guilford Press.

Grych, J. H., Wachsmuth-Schlaefer, T., & Klockow, L. L. (2002). Interparental aggression and young children's representations of family relationships. *Journal of Family Psychology, 16*(3), 259-272.

Gustafsson, B. M., Gustafsson, P. A., & Proczkowska-Björklund, M. (2016). The Strengths and Difficulties Questionnaire (SDQ) for preschool children - a Swedish validation. *Nordic journal of psychiatry, 70*(8), 567-574.

Gutman, L. M., Joshi, H., Khan, L., & Schoon, I. (2018). *Children of the millennium: Understanding the course of conduct problems during childhood*. London: Centre for Mental Health.

Halsey, L. G. (2019). The reign of the p-value is over: What alternative analyses could we employ to fill the power vacuum? *Biology Letters, 15*(5), 20190174.

Haltigan, J. D., Madigan, S., Bronfman, E., Bailey, H. N., Borland-Kerr, C., Mills-Koonce, R., & Lyons-Ruth, K. (2019). Refining the assessment of disrupted maternal communication: Using item response models to identify central indicators of disrupted behavior. *Development and Psychopathology, 31*(1), 261-277.

Haltigan, J. D., Roisman, G. I., & Fraley, R. C. (2013). The predictive significance of early caregiving experiences for symptoms of psychopathology through midadolescence: Enduring or transient effects? *Development and Psychopathology, 25*(1), 209-221.

Hannon, P. (2003). Developmental neuroscience: Implications for early childhood intervention and education. *Current Paediatrics, 13*(1), 58-63.

Harrison, L. J., & Murray, E. (2015). Stress, coping and wellbeing in kindergarten: Children's perspectives on personal, interpersonal and institutional challenges of school. *International Journal of Early Childhood, 47*(1), 79-103.

Harvey, E., Danforth, J. S., Ulaszek, W. R., & Eberhardt, T. L. (2001). Validity of the parenting scale for parents of children with attention-deficit/hyperactivity disorder. *Behaviour Research and Therapy, 39*(6), 731-743.

Haub, C., & Kaneda, T. (2013). *2013 World Population Data Sheet*. Washington, DC: Population Reference Bureau.

Hawes, D. J., & Dadds, M. R. (2006). Assessing parenting practices through parent-report and direct observation during parent-training. *Journal of Child and Family Studies, 15*(5), 554-567.

Hawkins, A. L., & Haskett, M. E. (2014). Internal working models and adjustment of physically abused children: The mediating role of self-regulatory abilities. *Journal of Child Psychology and Psychiatry, 55*(2), 135-143.

Hayes, A. F. (2018). *Introduction to mediation: A regression-based approach (methodology in the social sciences)* (2nd ed.). New York, NY: Guilford Press.

Hayes, N., & Bradley, S. (2009). *Right by children: Children's rights and rights based approaches to policy making in early childhood education and care: The case of Ireland*. Dublin, Ireland: Irish Research Council Humanities and the Social Sciences.

Heinrichs, N., Bertram, H., Kuschel, A., & Hahlweg, K. (2005). Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: Barriers to research and program participation. *Prevention Science, 6*(4), 275-286.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature, 466*(7302), 29-29.

Hesse, E. (2008). The Adult Attachment Interview: Protocol, method of analysis, and empirical studies. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 552-598). New York: The Guilford Press.

Hill, J., Fonagy, P., Lancaster, G., & Broyden, N. (2007). Aggression and intentionality in narrative responses to conflict and distress story stems: An investigation of boys with disruptive behaviour problems. *Attachment & human development, 9*(3), 223-237.

Hill, J., Murray, L., Leidecker, V., & Sharp, H. (2008). The dynamics of threat, fear and intentionality in the conduct disorders: Longitudinal findings in the children of women with post-natal depression. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1503), 2529-2541.

Hillman, S., Hodges, J., Steele, M., Cirasola, A., Asquith, K., & Kaniuk, J. (2020). Assessing changes in the internal worlds of early-and late-adopted children using the Story Stem Assessment Profile (SSAP). *Adoption & Fostering, 44*(4), 377-396.

Hinshaw, S. P., Han, S. S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21*(2), 143-150.

Hodges, J., & Steele, M. (2000). Effects of abuse on attachment representations: Narrative assessments of abused children. *Journal of Child Psychotherapy, 26*(3), 433-455.

Hodges, J., Steele, M., Hillman, S., Henderson, K., & Kaniuk, J. (2003). Changes in attachment representations over the first year of adoptive placement: Narratives of maltreated children. *Clinical Child Psychology and Psychiatry, 8*(3), 351-367.

Hoffman, M. L. (2001). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.

Holland, A. S., & Roisman, G. I. (2010). Adult attachment security and young adults' dating relationships over time: Self-reported, observational, and physiological evidence. *Developmental Psychology, 46*(2), 552-557.

Hood, A., & Waters, T. (2017). *Living standards, poverty and inequality in the UK: 2016-2017 to 2021-2022*. London: Institute for Fiscal Studies.

Hosny, N. M., Danquah, A., Berry, K., & Wan, M. W. (2020). Children's narratives of family life in Ghana: A cultural lens via story stems. *Journal of Child and Family Studies, 29*(12), 3521-3535.

Huaqing Qi, C., & Kaiser, A. P. (2003). Behavior problems of preschool children from low-income families: Review of the literature. *Topics in Early Childhood Special Education, 23*(4), 188-216.

Hubbs-Tait, L., Hughes, K. P., Culp, A. M., Osofsky, J. D., Hann, D. M., Eberhart-Wright, A., & Ware, L. M. (1996). Children of adolescent mothers: Attachment representation, maternal depression, and later behavior problems. *American Journal of Orthopsychiatry, 66*(3), 416-426.

Hutchings, J., Griffith, N., Bywater, T., & Williams, M. E. (2017). Evaluating the Incredible Years Toddler Parenting Programme with parents of toddlers in disadvantaged (Flying Start) areas of Wales. *Child: Care, Health and Development, 43*(1), 104-113.

Iles, J. E., Rosan, C., Wilkinson, E., & Ramchandani, P. G. (2017). Adapting and developing a video-feedback intervention for co-parents of infants at risk of externalising behaviour problems (VIPP-Co): A feasibility study. *Clinical Child Psychology and Psychiatry, 22*(3), 483-499.

Institute of Health Visiting. (2020a). *Health visiting in England: State of health visiting in England*. London: Institute of Health Visiting.

Institute of Health Visiting. (2020b). *State of health visiting in England: Are babies and their families being adequately supported in England in 2020 to get the best start in life?* . London: Institute of Health Visiting

Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Harder, V. S., Ang, R. P., Bilenberg, N., . . . Dias, P. (2010). Preschool psychopathology reported by parents in 23 societies: Testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(12), 1215-1224.

James, A. (2007). Giving voice to children's voices: Practices and problems, pitfalls and potentials. *American Anthropologist, 109*(2), 261-272.

James, A., Jenks, C., & Prout, A. (1998). Theorizing childhood. In C. Jenks (Ed.), *Childhood: Critical concepts in sociology*. Cambridge: Polity Press.

Janik Blaskova, L., & Gibson, J. L. (2021). Reviewing the link between language abilities and peer relations in children with developmental language disorder: The importance of children's own perspectives. *Autism & Developmental Language Impairments, 6*. doi:10.1177/23969415211021515

Johansson, E. (2011). Giving words to children's voices in research. In E. Johansoon & E. White (Eds.), *Educational research with our youngest: Voices of infants and toddlers* (pp. 1-14). New York, NY: Springer.

Jordan, S., Watkins, A., Storey, M., Allen, S. J., Brooks, C. J., Garaiova, I., . . . Russell, I. T. (2013). Volunteer bias in recruitment, retention, and blood sample donation in a randomised controlled trial involving mothers and their children at six months and two years: a longitudinal analysis. *PloS One, 8*(7), e67912.

Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2008a). *Manual VIPP-SD: Video-feedback program to promote positive parenting and sensitive discipline*. Netherlands: Leiden University.

Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2008b). *Promoting positive parenting: An attachment-based intervention*. Mahwah, NJ: Lawrence Erlbaum Associates.

Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2017a). Pairing attachment theory and social learning theory in video-feedback intervention to promote positive parenting. *Current Opinion in Psychology, 15*, 189-194.

Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2017b). Video-feedback intervention to promote positive parenting and sensitive discipline. In H. Steele & M. Steele (Eds.), *Handbook of attachment-based interventions* (pp. 1-26). New York: Guilford Press.

Juffer, F., Bakermans-Kranenburg, M., & van IJzendoorn, M. (2005). The importance of parenting in the development of disorganized attachment: Evidence from a preventive intervention study in adoptive families. *Journal of Child Psychology and Psychiatry, 46*(3), 263-274.

Juffer, F., & Steele, M. (2014). What words cannot say: The telling story of video in attachment-based interventions. *Attachment & human development, 16*(4), 307-314.

Juffer, F., Struis, E., Werner, C., & Bakermans-Kranenburg, M. (2017). Effective preventive interventions to support parents of young children: Illustrations from the Video-

feedback Intervention to promote Positive Parenting and Sensitive Discipline (VIPP-SD). *Journal of Prevention & Intervention in the Community, 45*(3), 202-214.

Kalinauskiene, L., Cekuoliene, D., van IJzendoorn, M., Bakermans-Kranenburg, M., Juffer, F., & Kusakovskaja, I. (2009). Supporting insensitive mothers: The Vilnius randomized control trial of video-feedback intervention to promote maternal sensitivity and infant attachment security. *Child: Care, Health and Development, 35*(5), 613-623.

Karlson, C. W., & Rapoff, M. A. (2009). Attrition in randomized controlled trials for pediatric chronic conditions. *Journal of Pediatric Psychology, 34*(7), 782-793.

Karreman, A., van Tuijl, C., van Aken, M. A., & Deković, M. (2006). Parenting and self-regulation in preschoolers: A meta-analysis. *Infant and Child Development, 15*(6), 561-579.

Keers, R., Coleman, J. R., Lester, K. J., Roberts, S., Breen, G., Thastum, M., . . . Eley, T. (2016). A genome-wide test of the differential susceptibility hypothesis reveals a genetic predictor of differential response to psychological treatments for child anxiety disorders. *Psychotherapy and Psychosomatics, 85*(3), 146-158.

Keiley, M. K., Lofthouse, N., Bates, J. E., Dodge, K. A., & Pettit, G. S. (2003). Differential risks of covarying and pure components in mother and teacher reports of externalizing and internalizing behavior across ages 5 to 14. *Journal of Abnormal Child Psychology, 31*(3), 267-283.

Keller, H., Bard, K., Morelli, G., Chaudhary, N., Vicedo, M., Rosabal-Coto, M., . . . Gottlieb, A. (2018). The myth of universal sensitive responsiveness: Comment on Mesman et al. (2017). *Child Development, 89*(5), 1921-1928.

Kessler, R. C., McLaughlin, K. A., Green, J. G., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., . . . Angermeyer, M. (2010). Childhood adversities and adult psychopathology in the WHO World Mental Health Surveys. *British Journal of Psychiatry, 197*(5), 378-385.

Kieling, C., & Rohde, L. A. (2012). Child and adolescent mental health research across the globe. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(9), 945-947.

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H., Milne, B. J., & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: Developmental follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry, 60*(7), 709-717.

Klein, A. M., Otto, Y., Fuchs, S., Zenger, M., & von Klitzing, K. (2013). Psychometric properties of the parent-rated SDQ in preschoolers. *European Journal of Psychological Assessment, 29*(2), 96-104.

Klein Velderman, M., Bakermans-Kranenburg, M., Juffer, F., & van IJzendoorn, M. (2006). Effects of attachment-based interventions on maternal sensitivity and infant attachment: Differential susceptibility of highly reactive infants. *Journal of Family Psychology, 20*(2), 266-274.

Kleine, D., Pearson, G., & Poveda, S. (2016). Method Guide 8. Participatory method: Engaging children's voices and experiences in research. Retrieved from http://globalkidsonline.net/wp-content/uploads/2016/05/Guide-8-Participatory-methods-Kleine-Pearson-Poveda.pdf

Knafo, A., & Plomin, R. (2006). Parental discipline and affection and children's prosocial behavior: Genetic and environmental links. *Journal of Personality and Social Psychology, 90*(1), 147-164.

Kochanska, G. (1997). Mutually responsive orientation between mothers and their young children: Implications for early socialization. *Child Development, 68*(1), 94-112.

Kochanska, G., & Aksan, N. (2006). Children's conscience and self-regulation. *Journal of Personality, 74*(6), 1587-1618.

Kochanska, G., Padavich, D. L., & Koenig, A. L. (1996). Children's narratives about hypothetical moral dilemmas and objective measures of their conscience: Mutual relations and socialization antecedents. *Child Development, 67*(4), 1420-1436.

Kok, R., Thijssen, S., Bakermans-Kranenburg, M., Jaddoe, V. W., Verhulst, F. C., White, T., . . . Tiemeier, H. (2015). Normal variation in early parental sensitivity predicts child structural brain development. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(10), 824-831.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent and teacher reports. *Journal of Child Psychology and Psychiatry, 34*(6), 991-1006.

Koller, D., & San Juan, V. (2015). Play-based interview methods for exploring young children's perspectives on inclusion. *International Journal of Qualitative Studies in Education, 28*(5), 610-631.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155-163.

Kopp, C. B. (1982). Antecedents of self-regulation: a developmental perspective. *Developmental Psychology, 18*(2), 199-214.

Köster, M., Ohmer, X., Nguyen, T. D., & Kärtner, J. (2016). Infants understand others' needs. *Psychological Science, 27*(4), 542-548.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry, 160*(9), 1566-1577.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241-253.

Kriss, A., Steele, H., & Steele, M. (2012). Measuring attachment and reflective functioning in early adolescence: An introduction to the friends and family interview. *Research in Psychotherapy: Psychopathology, Process and Outcome, 15*(2), 87-95.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606-613.

Kuriyan, A. B., Pelham, W. E., Molina, B. S., Waschbusch, D. A., Gnagy, E. M., Sibley, M. H., . . . Yu, J. (2013). Young adult educational and vocational outcomes of children diagnosed with ADHD. *Journal of Abnormal Child Psychology, 41*(1), 27-41.

Lahey, B. B., Loeber, R., Burke, J., Rathouz, P. J., & McBurnett, K. (2002). Waxing and waning in concert: Dynamic comorbidity of conduct disorder with other disruptive and emotional problems over 7 years among clinic-referred boys. *Journal of Abnormal Child Psychology 111*(4), 556-567.

Laible, D. (2007). Attachment with parents and peers in late adolescence: Links with emotional competence and social behavior. *Personality and Individual Differences, 43*(5), 1185-1197.

Laible, D., Carlo, G., Torquati, J., & Ontai, L. (2004). Children's perceptions of family relationships as assessed in a doll story completion task: Links to parenting, social competence, and externalizing behavior. *Social Development, 13*(4), 551-569.

Laible, D. J., & Thompson, R. A. (2002). Mother–child conflict in the toddler years: Lessons in emotion, morality, and relationships. *Child Development, 73*(4), 1187-1203.

Landry, S. H., & Smith, K. E. (2011). Maternal sensitivity and responsiveness: A conceptual framework with empirical evidence. In D. W. Davis & M. C. Logsdon (Eds.), *Maternal sensitivity: A scientific foundation for practice* (pp. 31-44). Hauppauge, NY: Nova Science Publishers.

Lansdown, G. (2001). *Promoting children's participation in democratic decision-making*. Florence, Italy: UNICEF Innocenti Research Centre.

Larzelere, R. E. (2000). Child outcomes of nonabusive and customary physical punishment by parents: An updated literature review. *Clinical Child and Family Psychology Review, 3*(4), 199-221.

Leavell, A. S., Tamis-LeMonda, C. S., Ruble, D. N., Zosuls, K. M., & Cabrera, N. J. (2012). African American, White and Latino fathers' activities with their sons and daughters in early childhood. *Sex Roles, 66*(1), 53-65.

Lilienfeld, S. O. (2003). Comorbidity between and within childhood externalizing and internalizing disorders: Reflections and directions. *Journal of Abnormal Child Psychology, 31*(3), 285-291.

Lincoln, C. R., Russell, B. S., Donohue, E. B., & Racine, L. E. (2017). Mother-Child interactions and preschoolers' emotion regulation outcomes: Nurturing autonomous emotion regulation. *Journal of Child and Family Studies, 26*(2), 559-573.

Lindsey, E. W., & Mize, J. (2001). Contextual differences in parent–child play: Implications for children's gender role development. *Sex Roles, 44*(3), 155-176.

Loheide-Niesmann, L., de Lijster, J., Hall, R., van Bakel, H., & Cima, M. (2021). Toddlers' preference for prosocial versus antisocial agents: No associations with empathy or attachment security. *Social Development, 30*(2), 410-427.

Longhi, E., Murray, L., Wellsted, D., Hunter, R., MacKenzie, K., Taylor-Colls, S., & Fearon, R. M. P. (2020). *Minding the baby® Home-visiting programme for vulnerable young mothers*. London: NSPCC.

Lorber, M. F., Xu, S., Slep, A. M. S., Bulling, L., & O'Leary, S. G. (2014). A new look at the psychometrics of the Parenting Scale through the lens of item response theory. *Journal of Clinical Child and Adolescent Psychology, 43*(4), 613-626.

Lorch, E. P., Milich, R., Flake, R. A., Ohlendorf, J., & Little, S. (2010). A developmental examination of story recall and coherence among children with ADHD. *Journal of Abnormal Child Psychology, 38*(3), 291-301.

Luby, J. L., Belden, A., Sullivan, J., & Spitznagel, E. (2007). Preschoolers' contribution to their diagnosis of depression and anxiety: Uses and limitations of young child self-report of symptoms. *Child Psychiatry and Human Development, 38*(4), 321-338.

Lucassen, N., Tiemeier, H., Luijk, M. P., Linting, M., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., . . . Lambregtse-Van den Berg, M. P. (2015). Expressed emotion during pregnancy predicts observed sensitivity of mothers and fathers in early childhood. *Parenting, 15*(3), 158-165.

Lundh, L. G., Wångby-Lundh, M., & Bjärehed, J. (2008). Self-reported emotional and behavioral problems in Swedish 14 to 15-year-old adolescents: A study with the self-report version of the strengths and difficulties questionnaire. *Scandinavian Journal of Psychology, 49*(6), 523-532.

Lundy, L. (2007). 'Voice' is not enough: Conceptualising Article 12 of the United Nations Convention on the Rights of the Child. *British Educational Research Journal, 33*(6), 927-942.

Lyons-Ruth, K., Todd Manly, J., von Klitzing, K., Tamminen, T., Emde, R., Fitzgerald, H., . . . Watanabe, H. (2017). The worldwide burden of infant mental and emotional disorder: Report of the task force of the world association for infant mental health. *Infant Mental Health Journal, 38*(6), 695-705.

Macfie, J., & Swan, S. A. (2009). Representations of the caregiver–child relationship and of the self, and emotion regulation in the narratives of young children whose mothers have borderline personality disorder. *Development and Psychopathology, 21*(3), 993-1011.

Mackenbach, J. D., Ringoot, A. P., Van Der Ende, J., Verhulst, F. C., Jaddoe, V. W., Hofman, A., . . . Tiemeier, H. W. (2014). Exploring the relation of harsh parental discipline with child emotional and behavioral problems by using multiple informants. The generation R study. *PloS One, 9*(8), e104793.

Madigan, S., Eirich, R., Racine, N., Borland-Kerr, C., Cooke, J. E., Devereux, C., . . . Haltigan, J. D. (2021). Feasibility of training service providers on the AMBIANCE-Brief measure for use in community settings. *Infant Mental Health Journal, 42*(3), 438-451.

Magnuson, K. A., & Duncan, G. J. (2002). Parents in poverty. In M. H. Bornstein (Ed.), *Handbook of parenting: Social conditions and applied parenting* (pp. 95-121). Mahwah, NJ: Lawrence Erlbaum.

Main, M., Goldwyn, R., & Hesse, E. (2002). *Adult attachment rating and classification systems*. Berkeley, CA: Unpublished manuscript.

Main, M., Kaplan, N., & Cassidy, J. (1985). Security in infancy, childhood, and adulthood: A move to the level of representation. In I. Bretherton & E. Waters (Eds.), *Growing points in attachment theory and research. Monographs of the Society for Research in Child Development* (Vol. 50, pp. 66-104). New York: Wiley.

Mandalia, D., Sadler, K., Vizard, T., Ford, T., Goodman, A., Goodman, R., & McManus, S. (2018). Mental health of children and young people in England, 2017: Behavioural disorders. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders, 42*(7), 1304-1313.

Maniglio, R. (2014). Prevalence of sexual abuse among children with conduct disorder: a systematic review. *Clinical Child and Family Psychology Review, 17*(3), 268-282.

Marmot, M. (2010). *Fair society, healthy lives: The Marmot review; strategic review of health inequalities in England post-2010*. London: Institute of Health Equity.

Marmot, M., & Allen, J. (2020). COVID-19: Exposing and amplifying inequalities. *Journal of Epidemiology and Community Health, 74*(9), 681-682.

Marmot, M., Allen, J., Boyce, T., Goldblatt, P., & Morrison, J. (2020). *Health equity in England: The Marmot review 10 years on*. London: Institute of Health Equity.

Masashi, N., & Tsutomu, N. (2012). The development of narrative coherence in preschoolers: Analysis of accounts about the experience of making cake. *Japanese Journal of Developmental Psychology, 23*(1), 66-74.

Maughan, B., Collishaw, S., Meltzer, H., & Goodman, R. (2008). Recent trends in UK child and adolescent mental health. *Social Psychiatry and Psychiatric Epidemiology, 43*(4), 305-310.

McCarthy, P. B., & Morote, E.-S. (2009). The link between investment in early childhood preshools and high school graduation rates for African American males in the United States of America. *Contemporary Issues in Early Childhood, 10*(3), 232-239.

McConaughy, S. H., Stanger, C., & Achenbach, T. M. (1992). Three-year course of behavioral/emotional problems in a national sample of 4- to 16-Year-Olds: I. Agreement among Informants. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*(5), 932-940.

McConnell, D., Breitkreuz, R., & Savage, A. (2011). From financial hardship to child difficulties: main and moderating effects of perceived social support. *Child: Care, Health and Development, 37*(5), 679-691.

McCulloch, A., Wiggins, R. D., Joshi, H. E., & Sachdev, D. (2000). Internalising and externalising children's behaviour problems in Britain and the US: Relationships to family resources 1. *Children & Society, 14*(5), 368-383.

McCullough, E., Stedmon, J., & Dallos, R. (2014). Narrative responses as an aid to understanding the presentation of maltreated children who meet criteria for autistic spectrum disorder and reactive attachment disorder: A case series study. *Clinical Child Psychology and Psychiatry, 19*(3), 392-411.

McDonald, A. M., Knight, R. C., Campbell, M. K., Entwistle, V. A., Grant, A. M., Cook, J. A., . . . Roberts, I. (2006). What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. *Trials, 7*(1), 1-9.

McKee, L., Roland, E., Coffelt, N., Olson, A. L., Forehand, R., Massari, C., . . . Zens, M. S. (2007). Harsh discipline and child problem behaviors: The roles of positive parenting and gender. *Journal of Family Violence, 22*(4), 187-196.

McLennan, J. D., & Offord, D. R. (2002). Should postpartum depression be targeted to improve child mental health? *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1), 28-35.

Measelle, J. R., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (1998). Assessing young children's views of their academic, social, and emotional lives: An evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Development, 69*(6), 1556-1576.

Meltzer, H., Gatward, R., Corbin, T., Goodman, R., & Ford, T. (2003a). *The mental health of young people looked after by local authorities in England*. London: The Stationery Office.

Meltzer, H., Gatward, R., Corbin, T., Goodman, R., & Ford, T. (2003b). *Persistence, onset, risk factors and outcomes of childhood mental disorders*. London: Office for National Statistics.

Menard, S. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage.

Menting, A. T., de Castro, B. O., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review, 33*(8), 901-913.

Mesman, J. (2021). Video observations of sensitivity in context: Integrating insights from seven cultural communities. *Attachment & human development, 23*(2), 212-221.

Mesman, J., & Emmen, R. A. (2013). Mary Ainsworth's legacy: A systematic review of observational instruments measuring parental sensitivity. *Attachment & human development, 15*(5-6), 485-506.

Mesman, J., Minter, T., Angnged, A., Cissé, I. A., Salali, G. D., & Migliano, A. B. (2018). Universality without uniformity: A culturally inclusive approach to sensitive responsiveness in infant caregiving. *Child Development, 89*(3), 837-850.

Mesman, J., Stolk, M., van Zeijl, J., Alink, L., Juffer, F., Bakermans-Kranenburg, M., . . . Koot, H. (2008). Extending the video-feedback intervention to sensitive discipline. *Promoting positive parenting: An attachment-based intervention (, pp. 171-192). New York, NY: Taylor and Francis*.

Mileva-Seitz, V. R., Ghassabian, A., Bakermans-Kranenburg, M., van den Brink, J. D., Linting, M., Jaddoe, V. W., . . . van IJzendoorn, M. (2015). Are boys more sensitive to sensitivity? Parenting and executive function in preschoolers. *Journal of Experimental Child Psychology, 130*, 193-208.

Miljkovitch, R., Pierrehumbert, B., Bretherton, I., & Halfon, O. (2004). Associations between parental and child attachment representations. *Attachment & human development, 6*(3), 305-325.

Miljkovitch, R., Pierrehumbert, B., & Halfon, O. (2007). Three-year-olds' attachment play narratives and their associations with internalizing problems. *Clinical Psychology & Psychotherapy, 14*(4), 249-257.

Miner, J. L., & Clarke-Stewart, K. A. (2008). Trajectories of externalizing behavior from age 2 to age 9: Relations with gender, temperament, ethnicity, parenting, and rater. *Developmental Psychology, 44*(3), 771-786.

Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PloS One, 13*(9), e0202855.

Minnis, H., Millward, R., Sinclair, C., Kennedy, E., Greig, A., Towlson, K., . . . Hill, J. (2006). The Computerized MacArthur Story Stem Battery – a pilot study of a novel medium for assessing children's representations of relationships. *International Journal of Methods in Psychiatric Research, 15*(4), 207-214.

Minze, L. C., McDonald, R., Rosentraub, E. L., & Jouriles, E. N. (2010). Making sense of family conflict: Intimate partner violence and preschoolers' externalizing problems. *Journal of Family Psychology, 24*(1), 5-11.

Mize, J., & Ladd, G. W. (1988). Predicting preschoolers' peer behavior and status from their interpersonal strategies: A comparison of verbal and enactive responses to hypothetical social dilemmas. *Developmental Psychology, 24*(6), 782-788.

Moran, G., Pederson, D. R., & Tarabulsy, G. M. (2011). Becoming sensitive to sensitivity: Lessons learned from the development of the Maternal Behavior Q-sort. In D. W. Davis & M. C. Logsdon (Eds.), *Maternal sensitivity: A scientific foundation for practice* (pp. 259-281). Hauppauge, NY: Nova Science Publishers.

Morrison, J., Pikhart, H., Ruiz, M., & Goldblatt, P. (2014). Systematic review of parenting interventions in European countries aiming to reduce social inequalities in children's health and development. *BMC Public Health, 14*(1), 1040. doi:10.1186/1471-2458-14-1040

Morsbach, S. K., & Prinz, R. J. (2006). Understanding and improving the validity of self-report of parenting. *Clinical Child and Family Psychology Review, 9*(1), 1-21.

Moss, E., Bureau, J.-F., Béliveau, M.-J., Zdebik, M., & Lépine, S. (2009). Links between children's attachment behavior at early school-age, their attachment-related representations, and behavior problems in middle childhood. *International Journal of Behavioral Development, 33*(2), 155-166.

Moss, E., Dubois-Comtois, K., Cyr, C., Tarabulsy, G. M., St-Laurent, D., & Bernier, A. (2011). Efficacy of a home-visiting intervention aimed at improving maternal sensitivity, child attachment, and behavioral outcomes for maltreated children: A randomized control trial. *Development and Psychopathology, 23*(1), 195-210.

Mueller, E., & Tingley, E. (1990). The Bears' Picnic: Children's representations of themselves and their families. *New Directions for Child and Adolescent Development, 1990*(48), 47-65.

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ). *European Child and Adolescent Psychiatry, 12*(1), 1-8.

Murphy, T. P., & Laible, D. J. (2013). The influence of attachment security on preschool children's empathic concern. *International Journal of Behavioral Development, 37*(5), 436-440.

Murray, E., & Harrison, L. J. (2005). Children's perspectives on their first year of school: Introducing a new pictorial measure of school stress. *European Early Childhood Education Research Journal, 13*(1), 111-127.

Murray, L., Woolgar, M., Briers, S., & Hipwell, A. (1999). Children's social representations in dolls' house play and theory of mind tasks, and their relation to family adversity and child disturbance. *Social Development, 8*(2), 179-200.

Myers, R. (1990). *Classical and modern regression with applications* (2nd ed.). Boston, MA: Duxbury.

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., O'Callaghan, M., . . . Shuttlewood, G. J. (2001). Bias influencing maternal reports of child behaviour and emotional state. *Social Psychiatry and Psychiatric Epidemiology, 36*(4), 186-194.

National Institute for Health and Care Excellence. (2013). *Antisocial behaviour and conduct disorders in children and young people: Recognition, intervention and management*. London: NICE.

Negrão, M., Pereira, M., Soares, I., & Mesman, J. (2014). Enhancing positive parent–child interactions and family functioning in a poverty sample: a randomized control trial. *Attachment & human development, 16*(4), 315-328.

Nelson, K. (1981). Social cognition in a script framework. In J. H. Flavell & L. Ross (Eds.), *Social cognitive development* (pp. 97-118). New York, NY: Cambridge University Press.

Niccols, A., & Feldman, M. (2006). Maternal sensitivity and behaviour problems in young children with developmental delay. *Infant and Child Development, 15*(5), 543-554.

NICHD Early Child Care Research Network. (2004). Trajectories of physical aggression from toddlerhood to middle childhood: Predictors, correlates, and outcomes. *Monographs of the Society for Research in Child Development, 69*(4), 1-129.

Nilsson, S., Björkman, B., Almqvist, A.-L., Almqvist, L., Björk-Willén, P., Donohue, D., . . . Hvit, S. (2015). Children's voices–differentiating a child perspective from a child's perspective. *Developmental Neurorehabilitation, 18*(3), 162-168.

Nowacki, K., Kliewer-Neumann, J., Bovenschen, I., Lang, K., Zimmermann, J., & Spangler, G. (2015). The relationship between attachment representations of foster parents and foster children and the role of the child's sex. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 64*(10), 733-751.

O'Brien, M., & Daley, D. (2011). Self-help parenting interventions for childhood behaviour disorders: A review of the evidence. *Child: Care, Health and Development, 37*(5), 623-637.

O'Connor, T. G., & Scott, S. (2007). *Parenting and outcomes in children* York: Joseph Rowntree Foundation.

O'Farrelly, C. (2021). Bringing young children's voices into programme development, randomized controlled trials and other unlikely places. *Children & Society, 35*(1), 34-47.

O'Farrelly, C., Barker, B., Watt, H., Babalis, D., Bakermans-Kranenburg, M., Byford, S., . . . Ramchandani, P. G. (2021). A video-feedback parenting intervention to prevent enduring behaviour problems in at-risk children aged 12-36 months: The Healthy Start, Happy Start RCT. *Health Technology Assessment, 25*(29), 1-84.

O'Farrelly, C., Booth, A., Tatlow-Golden, M., & Barker, B. (2020). Reconstructing readiness: Young children's priorities for their early school adjustment. *Early Childhood Research Quarterly, 50*, 3-16.

O'Farrelly, C., & Tatlow-Golden, M. (in press). It's up to you if you want to take part. Supporting young children's informed choice about research participation with simple visual booklets. *European Early Childhood Education Research Journal*.

O'Farrelly, C., Watt, H., Babalis, D., Bakermans-Kranenburg, M., Barker, B., Byford, S., . . . Ramchandani, P. G. (2021). A brief home-based parenting intervention to reduce behavior problems in young children: A pragmatic randomized clinical trial. *JAMA Pediatrics, 6*(175), 567-576.

O'Hara, L., Smith, E. R., Barlow, J., Livingstone, N., Herath, N. I., Wei, Y., . . . Macdonald, G. (2019). Video feedback for parental sensitivity and attachment security in children under five years. *Cochrane Database of Systematic Reviews*(11). doi:10.1002/14651858.CD012348.pub2.

O'Leary, S. G., & Vidair, H. B. (2005). Marital adjustment, child-rearing disagreements, and overreactive parenting: Predicting child behavior problems. *Journal of Family Psychology, 19*(2), 208-216.

Oakley, A., Strange, V., Bonell, C., Allen, E., & Stephenson, J. (2006). Process evaluation in randomised controlled trials of complex interventions. *BMJ, 332*(7538), 413-416.

Office for National Statistics. (2017). Families and households in the UK: 2017. Retrieved from https://www.ons.gov.uk/releases/familiesandhouseholdsintheuk2017

Office for National Statistics. (2019). Families and the labour market, UK: main dataset using the Labour Force Survey and Annual Population Survey. Retrieved from https://www.ons.gov.uk/releases/familiesandthelabourmarket

Office for National Statistics. (2020a). *Deaths involving COVID-19 by local area and socioeconomic deprivation: Deaths occurring between 1 March and 31 July 2020*. London: Office for National Statistics.

Office for National Statistics. (2020b). *Live births by sex, parity and ethnicity, England and Wales, 2014 to 2018 aggregated*. London: Office for National Statistics.

Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., & Lipman, E. L. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(8), 1078-1085.

Olds, D. L. (2006). The nurse–family partnership: An evidence-based preventive intervention. *Infant Mental Health Journal, 27*(1), 5-25.

Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D. W., . . . Holmberg, J. (2004). Effects of nurse home-visiting on maternal life course and child development: Age 6 follow-up results of a randomized trial. *Pediatrics, 114*(6), 1550-1559.

Oliver, B. R., & Pike, A. (2018). Mother-child positivity and negativity: Family-wide and child-specific main effects and interactions predict child adjustment. *Developmental Psychology, 54*(4), 744-756.

Oliver, B. R., & Pike, A. (2021). Introducing a novel online observation of parenting behavior: Reliability and validation. *Parenting, 21*(2), 168-183.

Opie, J. E., McIntosh, J. E., Esler, T. B., Duschinsky, R., George, C., Schore, A., . . . Olsson, C. A. (2020). Early childhood attachment stability and change: A meta-analysis. *Attachment & human development*, 1-34. doi:10.1080/14616734.2020.1800769

Oppenheim, D. (1997). The attachment doll-play interview for preschoolers. *International Journal of Behavioral Development, 20*(4), 681-697.

Oppenheim, D. (2006). Child, parent, and parent–child emotion narratives: Implications for developmental psychopathology. *Development and Psychopathology, 18*(3), 771-790.

Oppenheim, D., Emde, R., N, Hasson, M., & Warren, S. (1997). Preschoolers face moral dilemmas: A longitudinal study of acknowledging and resolving internal conflict. *International Journal of Psycho-Analysis, 78*, 943-957.

Oppenheim, D., Emde, R. N., & Warren, S. (1997). Children's narrative representations of mothers: Their development and associations with child and mother adaptation. *Child Development, 68*(1), 127-138.

Oppenheim, D., Nir, A., Warren, S., & Emde, R. N. (1997). Emotion regulation in mother–child narrative co-construction: Associations with children's narratives and adaptation. *Developmental Psychology, 33*(2), 284-294.

Ostrov, J. M., & Keating, C. F. (2004). Gender differences in preschool aggression during free play and structured interactions: An observational study. *Social Development, 13*(2), 255-277.

Page, T. (2001). The social meaning of children's narratives: A review of the attachment-based narrative story stem technique. *Child and Adolescent Social Work Journal, 18*(3), 171-187.

Page, T., Boris, N. W., Heller, S., Robinson, L., Hawkins, S., & Norwood, R. (2011). Narrative story stems with high risk six year-olds: Differential associations with mother-and teacher-reported psycho-social adjustment. *Attachment & human development, 13*(4), 359-380.

Page, T., & Bretherton, I. (2001). Mother- and father-child attachment themes in the story completions of pre-schoolers from post-divorce families: Do they predict relationships with peers and teachers? *Attachment & human development, 3*(1), 1-29.

Page, T., & Bretherton, I. (2003). Gender differences in stories of violence and caring by preschool children in post-divorce families: Implications for social competence. *Child and Adolescent Social Work Journal, 20*(6), 485-508.

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. F. (2014). Practitioner review: engaging fathers–recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry, 55*(11), 1187-1212.

Parent, J., McKee, L. G., & Forehand, R. (2016). Seesaw discipline: The interactive effect of harsh and lax discipline on youth psychological adjustment. *Journal of Child and Family Studies, 25*(2), 396-406.

Pascal, C., & Bertram, T. (2009). Listening to young citizens: The struggle to make real a participatory paradigm in research with young children. *European Early Childhood Education Research Journal, 17*(2), 249-262.

Pass, L., Arteche, A., Cooper, P., Creswell, C., & Murray, L. (2012). Doll play narratives about starting school in children of socially anxious mothers, and their relation to subsequent child school-based anxiety. *Journal of Abnormal Child Psychology, 40*(8), 1375-1384.

Patterson, G. R. (1982). *Coercive family process* (Vol. 3). Eugene, OR: Castalia Publishing Company.

Paulus, M., Licata, M., Kristen, S., Thoermer, C., Woodward, A., & Sodian, B. (2015). Social understanding and self-regulation predict pre-schoolers' sharing with friends and disliked peers: A longitudinal study. *International Journal of Behavioral Development, 39*(1), 53-64.

Pederson, D. R., Gleason, K. E., Moran, G., & Bento, S. (1998). Maternal attachment representations, maternal sensitivity, and the infant–mother attachment relationship. *Developmental Psychology, 34*(5), 925-933.

Pederson, D. R., Moran, G., & Bento, S. (1999). *Maternal behaviour Q-sort*. London, ON: University of Western Ontario.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology, 49*(12), 1373-1379.

Pereira, M., Negrão, M., Soares, I., & Mesman, J. (2014). Decreasing harsh discipline in mothers at risk for maltreatment: A randomized control trial. *Infant Mental Health Journal, 35*(6), 604-613.

Pesch, M. H., & Lumeng, J. C. (2017). Methodological considerations for observational coding of eating and feeding behaviors in children and their families. *International Journal of Behavioral Nutrition and Physical Activity, 14*(1), 1-14.

Peterson, C., & McCabe, A. (2004). Echoing our parents: Parental influences on children's narration. In M. W. Pratt & B. H. Fiese (Eds.), *Family Stories and the Life Course: Across Time and Generations* (pp. 27-53). Mahwah, NJ: Erlbaum.

Petitclerc, A., & Tremblay, R. E. (2009). Childhood disruptive behaviour disorders: Review of their origin, development, and prevention. *Canadian Journal of Psychiatry, 54*(4), 222-231.

Pickreign Stronach, E., Toth, S. L., Rogosch, F., Oshri, A., Manly, J. T., & Cicchetti, D. (2011). Child maltreatment, attachment security, and internal representations of mother and mother-child relationships. *Child Maltreatment, 16*(2), 137-145.

Pietromonaco, P. R., & Barrett, L. F. (2000). The internal working models concept: What do we really know about the self in relation to others? *Review of General Psychology, 4*(2), 155-175.

Pine, D. S., & Fox, N. A. (2015). Childhood antecedents and risk for adult mental disorders. *Annual Review of Psychology, 66*(3), 459-485.

Pingault, J.-B., Rijsdijk, F., Zheng, Y., Plomin, R., & Viding, E. (2015). Developmentally dynamic genome: Evidence of genetic influences on increases and decreases in conduct problems from early childhood to adolescence. *Scientific Reports, 5*. doi:10.1038/srep10053

Pinquart, M. (2017a). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology, 53*(5), 873-932.

Pinquart, M. (2017b). Associations of parenting dimensions and styles with internalizing symptoms in children and adolescents: A meta-analysis. *Marriage & Family Review, 53*(7), 613-640.

Pleck, J. H. (2012). Integrating father involvement in parenting research. *Parenting, 12*(2-3), 243-253.

Plokar, A., Bisaillon, C., & Terradas, M. (2018). Development of the Child Dissociation Assessment System using a narrative story stem task: A preliminary study. *European Journal of Trauma & Dissociation, 2*(1), 21-29.

Poehlmann, J. (2005). Representations of attachment relationships in children of incarcerated mothers. *Child Development, 76*(3), 679-696.

Portney, L. G. (2020). *Foundations of clinical research: Applications to evidence-based practice* (4th ed.). Philadelphia: FA Davis.

Poslawsky, I. E., Naber, F. B., Bakermans-Kranenburg, M., van Daalen, E., van Engeland, H., & van IJzendoorn, M. (2015). Video-feedback Intervention to promote Positive Parenting adapted to Autism (VIPP-AUTI): A randomized controlled trial. *Autism, 19*(5), 588-603.

Potegal, M., & Davidson, R. J. (2003). Temper tantrums in young children: 1. Behavioral composition. *Journal of Developmental and Behavioral Pediatrics, 24*(3), 140-147.

Power, T. G., & Chapieski, M. L. (1986). Childrearing and impulse control in toddlers: A naturalistic investigation. *Developmental Psychology, 22*(2), 271-275.

Prinzie, P., Onghena, P., & Hellinckx, W. (2007). Reexamining the Parenting Scale: Reliability, factor structure, and concurrent validity of a scale for assessing the discipline practices of mothers and fathers of elementary-school-aged children. *European Journal of Psychological Assessment, 23*(1), 24-31.

Pritchett, R., Rochat, T. J., Tomlinson, M., & Minnis, H. (2013). Risk factors for vulnerable youth in urban townships in South Africa: The potential contribution of reactive attachment disorder. *Vulnerable children and youth studies, 8*(4), 310-320.

Prout, A. (2011). Taking a step away from modernity: Reconsidering the new sociology of childhood. *Global Studies of Childhood, 1*(1), 4-14.

Prout, A., & James, A. (2015). A new paradigm for the sociology of childhood? Provenance, promise and problems. In A. James & A. Prout (Eds.), *Constructing and reconstructing childhood: Contemporary issues in the sociological study of childhood* (pp. 6-28). London: Routledge.

Public Health England. (2019). *Health visitor service delivery metrics (experimental statistics): 2018/2019 annual data statistical commentary (November 2019)*. London: Public Health England.

Pyle, A. (2013). Engaging young children in research through photo elicitation. *Early Child Development and Care, 183*(11), 1544-1558.

Quennerstedt, A., & Quennerstedt, M. (2014). Researching children's rights in education: Sociology of childhood encountering educational theory. *British Journal of Sociology of Education, 35*(1), 115-132.

Qvortrup, J. (2009). Are children human beings or human becomings?: A critical assessment of outcome thinking. *Rivista Internazionale di Scienze Sociali, 117*(3-4), 631-653.

Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (2015). The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development, 86*(3), 695-708.

Ramchandani, P. G., Murray, L., Romano, G., Vlachos, H., & Stein, A. (2011). An investigation of health anxiety in families where children have recurrent abdominal pain. *Journal of Pediatric Psychology, 36*(4), 409-419.

Ramchandani, P. G., O'Farrelly, C., Babalis, D., Bakermans-Kranenburg, M., Byford, S., Grimas, E., . . . Scott, S. (2017). Preventing enduring behavioural problems in young

children through early psychological intervention (Healthy Start, Happy Start): Study protocol for a randomized controlled trial. *Trials, 18*(1), 1-11.

Ramey, S. (2001). The science and art of parenting. Part II: Early influences o fparenting on achievement and competence. In J. Borkowski, S. Ramey, & M. Bristol-Power (Eds.), *Parenting and the child's world: Influences on academic, intellectual, and social-emotional development* (1st ed., pp. 47-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Rao, Z., Barker, B., O'Farrelly, C., & Ramchandani, P. G. (2021). Maternal anxiety and depression and their associations with mother–child pretend play: A longitudinal observational study. *BMC Psychology, 9*(1), 1-11.

Rao, Z., Fink, E., & Gibson, J. (2021). Dyadic association between aggressive pretend play and children's anger expression. *British Journal of Developmental Psychology, 39*(1), 153-168.

Raver, C. C., Carter, J. S., McCoy, D. C., Roy, A., Ursache, A., & Friedman, A. (2012). Testing models of children's self-regulation within educational contexts: Implications for measurement. In J. B. Benson (Ed.), *Advances in Child Development and Behavior* (Vol. 42, pp. 245-270). Waltham, MA: Elsevier.

Reese, E. (2008). Maternal coherence in the Adult Attachment Interview is linked to maternal reminiscing and to children's self concept. *Attachment & human development, 10*(4), 451-464.

Reese, E., Macfarlane, L., McAnally, H., Robertson, S.-J., & Taumoepeau, M. (2020). Coaching in maternal reminiscing with preschoolers leads to elaborative and coherent personal narratives in early adolescence. *Journal of Experimental Child Psychology, 189*, 104707. doi:10.1016/j.jecp.2019.104707

Reitman, D., Currier, R. O., Hupp, S. D., Rhode, P. C., Murphy, M. A., & O'Callaghan, P. M. (2001). Psychometric characteristics of the Parenting Scale in a Head Start population. *Journal of Clinical Child Psychology, 30*(4), 514-524.

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Harder, V. S., Otten, L., Bilenberg, N., . . . Dias, P. (2011). International comparisons of behavioral and emotional problems in preschool children: Parents' reports from 24 societies. *Journal of Clinical Child and Adolescent Psychology, 40*(3), 456-467.

Reynolds, C. R., & Kamphaus, R. W. (1998). *Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service.

Rhoades, K. A., & O'Leary, S. G. (2007). Factor structure and validity of the Parenting Scale. *Journal of Clinical Child and Adolescent Psychology, 36*(2), 137-146.

Richter, L. (2004). *The importance of caregiver-child interactions for the survival and healthy development of young children: A review*. Geneva, Switzerland: World Health Organisation.

Rinalidi, C. (2006). *In dialogue with Reggio Emilia: Listening, researching and learning*. Abingdon: Routledge.

Ringoot, A. P., Jansen, P. W., Rijlaarsdam, J., So, P., Jaddoe, V. W. V., Verhulst, F. C., & Tiemeier, H. (2017). Self-reported problem behavior in young children with and without a DSM-disorder in the general population. *European Psychiatry, 40*, 110-115.

Ringoot, A. P., Jansen, P. W., Steenweg-de Graaff, J., Measelle, J. R., van der Ende, J., Raat, H., . . . Tiemeier, H. (2013). Young children's self-reported emotional, behavioral, and peer problems: The Berkeley Puppet Interview. *Psychological Assessment, 25*(4), 1273-1285.

Ringoot, A. P., Tiemeier, H., Jaddoe, V. W., So, P., Hofman, A., Verhulst, F. C., & Jansen, P. W. (2015). Parental depression and child well-being: young children's self-reports helped addressing biases in parent reports. *Journal of Clinical Epidemiology, 68*(8), 928-938.

Roberts, H. (2000). Listening to children: And hearing them. In P. Christensen & A. James (Eds.), *Research with children* (pp. 225-241). London: Falmer Press.

Robinson, J., Hérot, C., Haynes, P., & Mantz-Simmons, L. (2000). Children's story stem responses: A measure of program impact on developmental risks associated with dysfunctional parenting. *Child Abuse and Neglect, 24*(1), 99-110.

Robinson, J., Mantz-Simmons, L., & Macfie, J. (1992). The MacArthur narrative coding manual. *Unpublished manuscript*.

Robinson, J. L. (2007). Story stem narratives with young children: Moving to clinical research and practice. *Attachment & human development, 9*(3), 179-185.

Robinson, L., Adair, P., Coffey, M., Harris, R., & Burnside, G. (2016). Identifying the participant characteristics that predict recruitment and retention of participants to randomised controlled trials involving children: a systematic review. *Trials, 17*(1). doi:10.1186/s13063-016-1415-0

Robling, M., Bekkers, M.-J., Bell, K., Butler, C. C., Cannings-John, R., Channon, S., . . . Kemp, A. (2016). Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): A pragmatic randomised controlled trial. *Lancet, 387*(10014), 146-155.

Robling, M., Lugg-Widger, F., Cannings-John, R., Sanders, J., Angel, L., Channon, S., . . . Slater, T. (2021). The Family Nurse Partnership to reduce maltreatment and improve child health and development in young children: The BB: 2–6 routine data-linkage follow-up to earlier RCT. *Public Health Research, 9*(2). doi:10.3310/phr09020

Rodrigues, M., Sokolovic, N., Madigan, S., Luo, Y., Silva, V., Misra, S., & Jenkins, J. (2021). Paternal sensitivity and children's cognitive and socioemotional outcomes: A meta-analytic review. *Child Development, 92*(2), 554-577.

Roisman, G. I., & Fraley, R. C. (2012). The legacy of early interpersonal experience. In J. B. Benson (Ed.), *Advances in Child Development and Behavior* (Vol. 42, pp. 79-112). Waltham, MA: Elsevier.

Ruzek, E., Jirout, J., Schenke, K., Vitiello, V., Whittaker, J. V., & Pianta, R. (2020). Using self report surveys to measure PreK children's academic orientations: A psychometric evaluation. *Early Childhood Research Quarterly, 50*, 55-66.

Sadler, K., Vizard, T., Ford, T., Marchesell, F., Pearce, N., Mandalia, D., . . . Goodman, A. (2018). Mental health of children and young people in England, 2017. Retrieved from https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017

Sadruddin, A. F. A., Ponguta, L. A., Zonderman, A. L., Wiley, K. S., Grimshaw, A., & Panter-Brick, C. (2019). How do grandparents influence child health and

development? A systematic review. *Social Science and Medicine, 239*, 112476. doi:10.1016/j.socscimed.2019.112476

Sainsbury Centre for Mental Health. (2009). *The chance of a lifetime: Preventing early conduct problems and reducing crime*. London: Sainsbury Centre for Mental Health London.

Sanson, A. V., Nicholson, J., Ungerer, J., Zubrick, S., Wilson, K., Ainley, J., . . . Harrison, L. (2002). *Introducing the longitudinal study of Australian children*. Melbourne, Australia: Australian Institute of Family Studies.

Sayal, K., Prasad, V., Daley, D., Ford, T., & Coghill, D. (2018). ADHD in children and young people: Prevalence, care pathways, and service provision. *The Lancet Psychiatry, 5*(2), 175-186.

Schechter, D. S., Zygmunt, A., Coates, S. W., Davies, M., Trabka, K. A., McCaw, J., . . . Robinson, J. L. (2007). Caregiver traumatization adversely impacts young children's mental representations on the MacArthur Story Stem Battery. *Attachment & human development, 9*(3), 187-205.

Schoeppe, S., Oliver, M., Badland, H. M., Burke, M., & Duncan, M. J. (2014). Recruitment and retention of children in behavioral health risk factor studies: REACH strategies. *International Journal of Behavioral Medicine, 21*(5), 794-803.

Scholtens, S., Rydell, A.-M., Bohlin, G., & Thorell, L. B. (2014). ADHD symptoms and attachment representations: considering the role of conduct problems, cognitive deficits and narrative responses in non-attachment-related story stems. *Journal of Abnormal Child Psychology, 42*(6), 1033-1042.

Scott, S., & Gardner, F. (2015). Parenting programs. In A. P. Thapar, DS, J. F. Leckman, M. Snowling, S. Scott, & E. Taylor (Eds.), *Rutter's child and adolescent psychiatry* (6th ed., pp. 483-496). Oxford: John Wiley & Sons.

Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: Follow up study of antisocial children into adulthood. *BMJ, 323*(7306), 191. doi:10.1136/bmj.323.7306.191

Semeniuk, Y. Y., & Riesch, S. K. (2011). Analysis of participant reactivity in dyads Performing a videotaped conflict-management task. *International Scholarly Research Notices*, 1-6. doi:10.5402/2011/596820

Shaw, D. S., & Bell, R. Q. (1993). Developmental theories of parental contributors to antisocial behavior. *Journal of Abnormal Child Psychology, 21*(5), 493-518.

Sher-Censor, E., & Oppenheim, D. (2004). Coherence and representations in preschoolers' narratives: Associations with attachment in infancy. In M. W. Pratt & B. H. Fiese (Eds.), *Family Stories and the Life Course: Across Time and Generations* (pp. 77-107). Mahwah, NJ: Erlbaum.

Shlafer, R. J., Raby, K. L., Lawler, J. M., Hesemeyer, P. S., & Roisman, G. I. (2015). Longitudinal associations between adult attachment states of mind and parenting quality. *Attachment & human development, 17*(1), 83-95.

Shmueli-Goetz, Y. (2014). The child attachment interview (CAI). In S. Farnfield & P. Holmes (Eds.), *The Routledge handbook of attachment: Assessment* (1st ed., pp. 133-146). London: Routledge.

Shmueli-Goetz, Y., Target, M., Fonagy, P., & Datta, A. (2008). The Child Attachment Interview: a psychometric study of reliability and discriminant validity. *Developmental Psychology, 44*(4), 939-956.

Sims, M., Guilfoyle, A., & Parry, T. S. (2006). Children's cortisol levels and quality of child care provision. *Child: Care, Health and Development, 32*(4), 453-466.

Skeen, S., Laurenzi, C. A., Gordon, S. L., du Toit, S., Tomlinson, M., Dua, T., . . . Servili, C. (2019). Adolescent mental health program components and behavior risk reduction: A meta-analysis. *Pediatrics, 144*(2), e20183488. doi:10.1542/peds.2018-3488

Skovgard, A. M. (2010). Mental health problems and psychopathology in infancy and early childhood. *Danish Medical Bulletin, 57*(10), 1-30.

Slade, A., Sadler, L., De Dios-Kenn, C., Webb, D., Currier-Ezepchick, J., & Mayes, L. (2005). Minding the baby: A reflective parenting program. *The Psychoanalytic Study of the Child, 60*(1), 74-100.

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly, 22*(2), 173-187.

Smith, G., Sylva, K., Smith, T., Sammons, P., & Omonigho, A. (2018). *Stop Start: Survival, decline or closure? Children's centres in England 2018*. London: The Sutton Trust.

Smyke, A. T., Zeanah, C. H., Fox, N. A., Nelson, C. A., & Guthrie, D. (2010). Placement in foster care enhances quality of attachment among young institutionalized children. *Child Development, 81*(1), 212-223.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. . In S. Leinhart (Ed.), *Sociological Methodology* (pp. 290-312). San Francisco, CA: Jossey-Bass.

Society for Emotion and Attachment Studies. (2021). Explanations of attachment theoretical concepts. Version April 2021. Retrieved from https://seasinternational.org/explanations-of-attachment-theoretical-concepts/

Socolar, R. R., & Stein, R. E. (1996). Maternal discipline of young children: Context, belief, and practice. *Journal of Developmental and Behavioral Pediatrics, 17*(1), 1-8.

Solomon, J., & George, C. (2008). The measurement of attachment security and related constructs in infancy and early childhood. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 383-416). New York: The Guilford Press.

Sonuga-Barke, E. J., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(4), 402-408.

Sonuga-Barke, E. J., Thompson, M., Abikoff, H., Klein, R., & Brotman, L. M. (2006). Nonpharmacological interventions for preschoolers with ADHD: The case for specialized parent training. *Infants & Young Children, 19*(2), 142-153.

Sonuga-Barke, E. J. S., Barton, J., Daley, D., Hutchings, J., Maishman, T., Raftery, J., . . . Thompson, M. J. J. (2018). A comparison of the clinical effectiveness and cost of specialised individually delivered parent training for preschool attention-deficit/hyperactivity disorder and a generic, group-based programme: a multi-centre,

randomised controlled trial of the New Forest Parenting Programme versus Incredible Years. *European Child and Adolescent Psychiatry, 27*(6), 797-809. doi:10.1007/s00787-017-1054-3

Sonuga-Barke, E. J., Lamparelli, M., Stevenson, J., Thompson, M., & Henry, A. (1994). Behaviour problems and pre-school intellectual attainment: The associations of hyperactivity and conduct problems. *Journal of Child Psychology and Psychiatry, 35*(5), 949-960.

Sonuga-Barke, E. J., Thompson, M., Daley, D., & Laver-Bradbury, C. (2004). Parent training for attention deficit/hyperactivity disorder: Is it as effective when delivered as routine rather than as specialist care? *British Journal of Clinical Psychology, 43*(4), 449-457.

Spencer, N., Devereux, E., Wallace, A., Sundrum, R., Shenoy, M., Bacchus, C., & Logan, S. (2005). Disabling conditions and registration for child abuse and neglect: A population-based study. *Pediatrics, 116*(3), 609-613.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092-1097.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA, 282*(18), 1737-1744.

Spitzer, R. L., Williams, J. B., Kroenke, K., Hornyak, R., McMurray, J., & Group, P. H. Q. O.-G. S. (2000). Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *American Journal of Obstetrics and Gynecology, 183*(3), 759-769.

Sroufe, L. A., Coffino, B., & Carlson, E. A. (2010). Conceptualizing the role of early experience: Lessons from the Minnesota longitudinal study. *Developmental Review, 30*(1), 36-51.

Stacks, A. M. (2007). Defensive dysregulation in preschool children's attachment story narratives and its relation to attachment classification and externalizing behaviour. *School Psychology International, 28*(3), 294-312.

Stams, G.-J. J., Juffer, F., Rispens, J., & Hoksbergen, R. A. (2000). The development and adjustment of 7-year-old children adopted in infancy. *Journal of Child Psychology and Psychiatry, 41*(8), 1025-1037.

Stams, G.-J. J., Juffer, F., & van IJzendoorn, M. (2002). Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Developmental Psychology, 38*(5), 806-821.

Stanley, T. D., Carter, E. C., & Doucouliagos, H. (2018). What meta-analyses reveal about the replicability of psychological research. *Psychological Bulletin, 144*(12), 1325-1346.

Steele, H., Bate, J., Steele, M., Dube, S. R., Danskin, K., Knafo, H., . . . Murphy, A. (2016). Adverse childhood experiences, poverty, and parenting stress. *Canadian Journal of Behavioural Science, 48*(1), 32-38.

Steele, H., & Steele, M. (2005). The construct of coherence as an indicator of attachment security in middle childhood: The Friends and Family Interview. In K. A. Kerns & R. A. Richardson (Eds.), *Attachment in middle childhood* (pp. 137-160). New York, NY: The Guilford Press.

Steele, R. D., Waters, T. E., Bost, K. K., Vaughn, B. E., Truitt, W., Waters, H. S., . . . Roisman, G. I. (2014). Caregiving antecedents of secure base script knowledge: A comparative analysis of young adult attachment representations. *Developmental Psychology, 50*(11), 2526-2538.

Steele, R. G., Nesbitt-Daly, J. S., Daniel, R. C., & Forehand, R. (2005). Factor structure of the Parenting Scale in a low-income African American sample. *Journal of Child and Family Studies, 14*(4), 535-549.

Stephenson, A. (2009). Horses in the sandpit: Photography, prolonged involvement and 'stepping back' as strategies for listening to children's voices. *Early Child Development and Care, 179*(2), 131-141. doi:10.1080/03004430802667047

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NJ: Erlbaum.

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child Development, 81*(6), 1814-1827.

Tabachnick, B., & Fidell, L. (2014). *Using multivariate statistics* (6th ed.). Essex: Pearson Education.

Tamis-LeMonda, C. S., & Baumwell, L. (2011). Parental sensitivity in early development: Definition, methods, measurement, and generalizability. In D. Winders Davis & M. C. Logsdon (Eds.), *Maternal sensitivity: A scientific foundation for practice* (pp. 1-17). Hauppauge, NY: Nova Science Publishers.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748-767.

Tamis-LeMonda, C. S., Bornstein, M. H., Baumwell, L., & Melstein Damast, A. (1996). Responsive parenting in the second year: Specific influences on children's language and play. *Early Development and Parenting, 5*(4), 173-183.

Tangen, R. (2008). Listening to children's voices in educational research: Some theoretical and methodological problems. *European Journal of Special Needs Education, 23*(2), 157-166.

Tatlow-Golden, M., & Montgomery, H. (2021). Childhood Studies and child psychology: Disciplines in dialogue? *Children & Society, 35*(1), 3-17. doi:10.1111/chso.12384

Tay-Lim, J., & Lim, S. (2013). Privileging younger children's voices in research: Use of drawings and a co-construction process. *International Journal of Qualitative Methods, 12*(1), 65-83.

Taylor, E., Everitt, B., Rutter, M., Thorley, G., Schachar, R., & Wieselberg, M. (1986a). Conduct disorder and hyperactivity: I: Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *British Journal of Psychiatry, 149*(6), 760-767.

Taylor, E., Everitt, B., Rutter, M., Thorley, G., Schachar, R., & Wieselberg, M. (1986b). Conduct disorder and hyperactivity: II: A cluster analytic approach to the

identification of a behavioural syndrome. *British Journal of Psychiatry, 149*(6), 768-777.

Taylor, E., Sandberg, S., Thorley, G., & Giles, S. (1991). *The epidemiology of childhood hyperactivity*. Oxford: Oxford University Press.

Teti, D. M., Cole, P. M., Cabrera, N., Goodman, S. H., & McLoyd, V. C. (2017). Supporting parents: How six decades of parenting research can inform policy and best practice. *Social Policy Report, 30*(5), 164-176.

Teubert, D., & Pinquart, M. (2010). The association between coparenting and child adjustment: A meta-analysis. *Parenting: Science and Practice, 10*(4), 286-307.

Theunissen, M. H., Vogels, A. G., de Wolff, M. S., & Reijneveld, S. A. (2013). Characteristics of the strengths and difficulties questionnaire in preschool children. *Pediatrics, 131*(2), e446-e454.

Thompson, R. A. (2007). The development of the person: Social understanding, relationships, conscience, self. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (6th ed., Vol. 3, pp. 24-98). New York: John Wiley & Sons.

Thompson, R. A. (2015). Early attachment and later development: Familiar questions, new answers. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment* (2nd ed., pp. 348-365). New York, NY: Guilford Press.

Thompson, R. A. (2018). Early attachment and later development: Reframing the questions. In J. Cassidy & P. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (3rd ed., pp. 348-365). New York: Guilford Press.

Tomlinson, C., Bond, C., & Hebron, J. (2020). The school experiences of autistic girls and adolescents: A systematic review. *European Journal of Special Needs Education, 35*(2), 203-219.

Tomlinson, M., Bornstein, M. H., Marlow, M., & Swartz, L. (2014). Imbalances in the knowledge about infant mental health in rich and poor countries: Too little progress in bridging the gap. *Infant Mental Health Journal, 35*(6), 624-629.

Torng, P.-C., & Sah, W.-H. (2020). Narrative abilities of Mandarin-speaking children with and without specific language impairment: Macrostructure and microstructure. *Clinical Linguistics & Phonetics, 34*(5), 453-478.

Toroyan, T., Oakley, A., Laing, G., Roberts, I., Mugford, M., & Turner, J. (2004). The impact of day care on socially disadvantaged families: an example of the use of process evaluation within a randomized controlled trial. *Child: Care, Health and Development, 30*(6), 691-698.

Toth, S. L., Cicchetti, D., Macfie, J., & Emde, R. N. (1997). Representations of self and other in the narratives of neglected, physically abused, and sexually abused preschoolers. *Development and Psychopathology, 9*(4), 781-796.

Toth, S. L., Maughan, A., Manly, J. T., Spagnola, M., & Cicchetti, D. (2002). The relative efficacy of two interventions in altering maltreated preschool children's representational models: Implications for attachment theory. *Development and Psychopathology, 14*(4), 877-908.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

Tremblay, R. E., Nagin, D. S., Séguin, J. R., Zoccolillo, M., Zelazo, P. D., Boivin, M., . . . Japel, C. (2004). Physical aggression during early childhood: Trajectories and predictors. *Pediatrics, 114*(1), e43-e50.

Truong, S., & Mahon, M. (2012). Through the lens of participatory photography: Engaging Thai children in research about their community play centre. *International Journal of Play, 1*(1), 75-90.

United Nations Committee on the Rights of the Child. (2005). *General comment No. 7 (2005): Implementing child rights in early childhood*. Geneva: United Nations.

United Nations Convention on the Rights of the Child. (1989). New York: United Nations.

Urbina-Garcia, M. A. (2019). Methodological strategies to listen to children's voices: A systematic critical review. *Revista Colombiana de Educación, 1*(77), 61-85.

Valla, J.-P., Bergeron, L., & Smolla, N. (2000). The Dominic-R: A pictorial interview for 6-to 11-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(1), 85-93.

van Daal, J., Verhoeven, L., & van Balkom, H. (2007). Behaviour problems in children with language impairment. *Journal of Child Psychology and Psychiatry, 48*(11), 1139-1147.

van Der Voort, A., Juffer, F., & Bakermans-Kranenburg, M. (2014). Sensitive parenting is the foundation for secure attachment relationships and positive social-emotional development of children. *Journal of Children's Services, 9*(2), 165-176.

van IJzendoorn, M. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin, 117*(3), 387-403.

van IJzendoorn, M., & Bakermans-Kranenburg, M. (2012). Integrating temperament and attachment: The differential susceptibility paradigm. In M. Zentner & R. L. Shiner (Eds.), *Handbook of temperament* (pp. 403-424). New York, NY: Guilford Press.

van Zeijl, J., Mesman, J., Stolk, M. N., Alink, L. R., van IJzendoorn, M., Bakermans-Kranenburg, M., . . . Koot, H. M. (2006). Terrible ones? Assessment of externalizing behaviors in infancy with the Child Behavior Checklist. *Journal of Child Psychology and Psychiatry, 47*(8), 801-810.

Vasileva, M., Graf, R. K., Reinelt, T., Petermann, U., & Petermann, F. (2021). Research review: A meta-analysis of the international prevalence and comorbidity of mental disorders in children between 1 and 7 years. *Journal of Child Psychology and Psychiatry, 62*(4), 372-381.

Veale, A. (2005). Creative methodologies in participatory research with children. In S. Greene & D. Hogan (Eds.), *Researching children's experience: Approaches and methods* (pp. 253-272). London: Sage.

Verhage, M. L., Schuengel, C., Madigan, S., Fearon, R. M. P., Oosterman, M., Cassibba, R., . . . van IJzendoorn, M. (2016). Narrowing the transmission gap: A synthesis of three decades of research on intergenerational transmission of attachment. *Psychological Bulletin, 142*(4), 337-366.

von Klitzing, K., & Bürgin, D. (2005). Parental capacities for triadic relationships during pregnancy: Early predictors of children's behavioral and representational functioning at preschool age. *Infant Mental Health Journal, 26*(1), 19-39.

von Klitzing, K., Döhnert, M., Kroll, M., & Grube, M. (2015). Mental disorders in early childhood. *Deutsches Arzteblatt International, 112*(21-22), 375-386.

von Klitzing, K., Kelsay, K., Emde, R. N., Robinson, J., & Schmitz, S. (2000). Gender-specific characteristics of 5-year-olds' play narratives and associations with behavior ratings. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*(8), 1017-1023.

von Klitzing, K., Stadelmann, S., & Perren, S. (2007). Story stem narratives of clinical and normal kindergarten children: Are content and performance associated with children's social competence? *Attachment & human development, 9*(3), 271-286.

Vygotsky, L. S. (1978). Socio-cultural theory. In M. Cole (Ed.), *Mind in society* (pp. 52-58). Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1987). *The collected works of LS Vygotsky: Vol. 2.The fundamentals of defectology*. New York: Plenum.

Wakschlag, L. S., Choi, S. W., Carter, A. S., Hullsiek, H., Burns, J., McCarthy, K., . . . Briggs-Gowan, M. J. (2012). Defining the developmental parameters of temper loss in early childhood: implications for developmental psychopathology. *Journal of Child Psychology and Psychiatry, 53*(11), 1099-1108.

Wakschlag, L. S., Roberts, M. Y., Flynn, R. M., Smith, J. D., Krogh-Jespersen, S., Kaat, A. J., . . . Norton, E. S. (2019). Future directions for early childhood prevention of mental disorders: A road map to mental health, earlier. *Journal of Clinical Child and Adolescent Psychology, 48*(3), 539-554.

Wakschlag, L. S., Tolan, P. H., & Leventhal, B. L. (2010). 'Ain't misbehavin': Towards a developmentally-specified nosology for preschool disruptive behavior. *Journal of Child Psychology and Psychiatry, 51*(1), 3-22.

Walker, J. L., Lahey, B. B., Russo, M. F., Frick, P. J., Christ, M. A. G., McBurnett, K., . . . Green, S. M. (1991). Anxiety, inhibition, and conduct disorder in children: I. Relations to social impairment. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(2), 187-191.

Wan, M. W., & Green, J. (2010). Negative and atypical story content themes depicted by children with behaviour problems. *Journal of Child Psychology and Psychiatry, 51*(10), 1125-1131.

Warren, S. L., Oppenheim, D., & Emde, R. N. (1996). Can emotions and themes in children's play predict behavior problems? *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(10), 1331-1337.

Waters, T. E., Raby, K. L., Ruiz, S. K., Martin, J., & Roisman, G. I. (2018). Adult attachment representations and the quality of romantic and parent–child relationships: An examination of the contributions of coherence of discourse and secure base script knowledge. *Developmental Psychology, 54*(12), 2371-2381.

Weber, L., Kamp-Becker, I., Christiansen, H., & Mingebach, T. (2019). Treatment of child externalizing behavior problems: A comprehensive review and meta–meta-analysis on effects of parent-based interventions on parental characteristics. *European Child and Adolescent Psychiatry, 28*(8), 1025-1036.

Webster-Stratton, C. (1985). Comparisons of behavior transactions between conduct-disordered children and their mothers in the clinic and at home. *Journal of Abnormal Child Psychology, 13*(2), 169-183.

Webster-Stratton, C. (2001). The Incredible Years: Parents, teachers, and children training series. *Residential Treatment for Children & Youth, 18*(3), 31-45.

Webster-Stratton, C. (2008). *The Incredible Years Parents, Babies and Toddlers series. Leader guide manual*. Seattle, WA: Incredible Years Limited.

Weiss, B., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1992). Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. *Child Development, 63*(6), 1321-1335.

Werner, C. D., Vermeer, H. J., Linting, M., & van IJzendoorn, M. (2018). Video-feedback intervention in center-based child care: A randomized controlled trial. *Early Childhood Research Quarterly, 42*, 93-104.

Wichstrøm, L., Berg-Nielsen, T. S., Angold, A., Egger, H. L., Solheim, E., & Sveen, T. H. (2012). Prevalence of psychiatric disorders in preschoolers. *Journal of Child Psychology and Psychiatry, 53*(6), 695-705.

Willoughby, M., Kupersmidt, J., Voegler-Lee, M., & Bryant, D. (2011). Contributions of hot and cool self-regulation to preschool disruptive behavior and academic achievement. *Developmental Neuropsychology, 36*(2), 162-180.

Windhorst, D. A., Mileva-Seitz, V. R., Linting, M., Hofman, A., Jaddoe, V. W., Verhulst, F. C., . . . Bakermans-Kranenburg, M. (2015). Differential susceptibility in a developmental perspective: DRD4 and maternal sensitivity predicting externalizing behavior. *Developmental Psychobiology, 57*(1), 35-49.

Winstanley, A., & Gattis, M. (2013). The Baby Care Questionnaire: A measure of parenting principles and practices during infancy. *Infant Behavior and Development, 36*(4), 762-775.

Winston, R., & Chicot, R. (2016). The importance of early bonding on the long-term mental health and resilience of children. *London Journal of Primary Care, 8*(1), 12-14.

Wolcott, C. S., Williford, A. P., & Hartz-Mandell, K. (2019). The validity of narratives for understanding children's perceptions of the teacher–child relationship for preschoolers who display elevated disruptive behaviors. *Early Education and Development, 30*(7), 887-912.

Wong, M., Bost, K. K., Shin, N., Veríssomo, M., Maia, J., Monteiro, L., . . . Vaughn, B. E. (2011). Preschool children's mental representations of attachment: Antecedents in their secure base behaviors and maternal attachment scripts. *Attachment & human development, 13*(5), 489-502.

Woolgar, M. (1999). Projective doll play methodologies for preschool children. *Child Psychology and Psychiatry Review, 4*(3), 126-134.

Woolgar, M., & Murray, L. (2010). The representation of fathers by children of depressed mothers: refining the meaning of parentification in high-risk samples. *Journal of Child Psychology and Psychiatry, 51*(5), 621-629.

Woolgar, M., Steele, H., Steele, M., Yabsley, S., & Fonagy, P. (2001). Children's play narrative responses to hypothetical dilemmas and their awareness of moral emotions. *British Journal of Developmental Psychology, 19*(1), 115-128.

Woolley, H., Hertzmann, L., & Stein, A. (2007). Video-feedback intervention with mothers with postnatal eating disorders and their infants. In F. Juffer, M. Bakermans-Kranenburg, & M. van IJzendoorn (Eds.), *Promoting positive parenting: An attachment based intervention* (pp. 131-158). New York, NY: Erlbaum Associates.

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. (2009). Measuring adolescent psychopathology: Psychometric properties of the self-report Strengths and Difficulties Questionnaire in a sample of Chinese adolescents. *Journal of Adolescent Health, 45*(1), 55-62.

Yuval-Adler, S., & Oppenheim, D. (2015). Story completion play narrative methods for preschool children. In O. Saracho (Ed.), *Handbook of research methods in early childhood education: Review of research methodologies* (pp. 323-381). University of Maryland, US: Information Age Publishing.

Zahn-Waxler, C., Cole, P. M., Richardson, D. T., Friedman, R. J., Michel, M. K., & Belouad, F. (1994). Social problem solving in disruptive preschool children: Reactions to hypothetical situations of conflict and distress. *Merrill-Palmer Quarterly, 40*(1), 98-119.

Zahn-Waxler, C., Park, J.-H., Usher, B., Belouad, F., Cole, P., & Gruber, R. (2008). Young children's representations of conflict and distress: A longitudinal study of boys and girls with disruptive behavior problems. *Development and Psychopathology, 20*(1), 99-119.

Zahn-Waxler, C., Friedman, R. J., Cole, P. M., Mizuta, I., & Hiruma, N. (1996). Japanese and United States preschool children's responses to conflict and distress. *Child Development, 67*(5), 2462-2477.

Zaslow, M. J., Weinfield, N. S., Gallagher, M., Hair, E. C., Ogawa, J. R., Egeland, B., . . . De Temple, J. M. (2006). Longitudinal prediction of child outcomes from differing measures of parenting in a low-income sample. *Developmental Psychology, 42*(1), 27-37.

Zeanah Jr, C. H., & Zeanah, P. D. (2019). Infant mental health: The clinical science of early experience. In C. H. Zeanah (Ed.), *Handbook of infant mental health* (4th ed., pp. 5-24). New York: Guilford Press.

Zenaro, M. P., Rossi, N. F., de Souza, A. L. D. M., & Giacheti, C. M. (2019). Oral narrative structure and coherence of children with attention deficit hyperactivity disorder. *CoDAS, 31*(6), 1-8. doi:10.1590/2317-1782/20192018197

Zimmerman, D. W. (2004). A note on preliminary tests of equality of variances. *British Journal of Mathematical and Statistical Psychology, 57*(1), 173-181.

Appendices

Appendix A

Ethical Approvals for the Healthy Start, Happy Start Study

NHS
Health Research Authority

NRES Committee London - Riverside
Level 3 Block B
Whitefriars
Lewins Mead
Bristol
BS1 2NT

[redacted]

12 December 2014

Dr Paul Ramchandani
Reader in Child and Adolescent Psychiatry and Consultant Child and Adolescent Psychiatrist
Imperial College London / CNWL NHS Foundation Trust
Centre for Mental Health
Commonwealth Building, Hammersmith Hospital
London
W12 0NN

Dear Dr Ramchandani

Study title:	**Preventing enduring behavioural problems in young children through early psychological intervention: Healthy Start, Happy Start.**
REC reference:	**14/LO/2071**
Protocol number:	**14HH2370**
IRAS project ID:	**160786**

The Research Ethics Committee reviewed the above application at the meeting held on 01 December 2014. Thank you for attending to discuss the application.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager Miss [redacted]

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below. .

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

1) **Changes to the Participant Information Sheet (PIS)**

a) The Committee requested that the PIS states clearly what will happen if disclosures that cause concerns for the patient safety or information about criminal activity will be reported and confidentiality breached.

b) The PIS should state clearly that the questionnaires used in the study are currently used in standard care.

2) **Changes to the Consent form**

a) Please include a sentence to state that the participant's personal data will be stored and may be used for future related research.

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. The REC will acknowledge receipt and provide a final list of the approved documentation for the study, which can be made available to host organisations to facilitate their permission for the study. Failure to provide the final versions to the REC may cause delay in obtaining permissions.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on question 2 of the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact [redacted], the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non NHS sites

The Committee has not yet completed any site-specific assessment(s) (SSA) for the non-NHS research site(s) taking part in this study. The favourable opinion does not therefore apply to any non-NHS site at present. I will write to you again as soon as an SSA application(s) has been reviewed. In the meantime no study procedures should be initiated at non-NHS sites.

Summary of discussion at the meeting

The Committee stated that they were encouraged to see that they had developed their pilot study into this second stage trial. The Committee asked you to clarify what arrangements are in place to safeguard children taking part in the study and if confidentiality will be breached if information given compromised the participant's safety.

You explained that the research nurses will report to the Chief Investigator if there were concerns about the participant's safety. The clinicians delivering the interventions work for the NHS and comply with standard practice for reporting disclosures and supporting patients appropriately. You added that the participant will be put into contact with local support services.

The Committee requested that the PIS states clearly what will happen if disclosures that cause concerns for the patient safety or information about criminal activity will be reported and confidentiality breached.

The Committee asked you if travel expenses will be reimbursed.

You replied that you do not anticipate participants making trips as most of the study interventions are held at the participant's home.

The Committee thanked you for attending the meeting and you left the room.

Other ethical issues were raised and resolved in preliminary discussion before your attendance at the meeting.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Copies of advertisement materials for research participants [Poster A]	1.0	04 November 2014
Copies of advertisement materials for research participants [Poster B]	1.0	04 November 2014
Covering letter on headed paper [Cover Letter]		04 November 2014
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Insurance]	N/A	21 July 2014
GP/consultant information sheets or letters [GP Letter]	1.0	04 November 2014
GP/consultant information sheets or letters [Health Visitor Letter]	1.0	04 November 2014
IRAS Checklist XML [Checklist_05112014]		05 November 2014
Letter from sponsor [Sponsor Letter]	N/A	04 November 2014
Letter from statistician [Letter from Statistician]	N/A	31 October 2014
Letters of invitation to participant [Invitation Letter]	1.0	04 November 2014
Non-validated questionnaire [Feedback Questionnaire]	1.0	04 November 2014
Other [Thank you letter - no inclusion]	1.0	04 November 2014
Participant consent form [Consent Form Phase 1]	1.0	04 November 2014
Participant consent form [Consent Form Phase 2]	1.0	04 November 2014
Participant information sheet (PIS) [Participant Information Sheet Phase 1]	1.0	04 November 2014
Participant information sheet (PIS) [Participant Information Sheet Phase 1]	1.0	04 November 2014
REC Application Form [REC_Form_05112014]		05 November 2014
Research protocol or project proposal [Trial Protocol]	1.0	04 November 2014
Summary CV for Chief Investigator (CI) [Chief Investigator CV]	N/A	04 October 2014
Validated questionnaire [Audit-C questionnaire]	1.0	04 November 2014
Validated questionnaire [BDI questionnaire]	1.0	04 November 2014
Validated questionnaire [CBCL questionnaire]	1.0	04 November 2014
Validated questionnaire [DAS questionnaire]	1.0	04 November 2014
Validated questionnaire [GAD-7 questionnaire]	1.0	04 November 2014
Validated questionnaire [Parenting scale]	1.0	04 November 2014
Validated questionnaire [Demographics questionnaire]	1.0	04 November 2014
Validated questionnaire [Pre-PACS]	1.0	04 November 2014
Validated questionnaire [SDQ screening questionnaire]	1.0	04 November 2014
Validated questionnaire [SDQ Full]		
Validated questionnaire [CA-SUS]	1.0	04 November 2014

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

There were no declarations of interest.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

14/LO/2071	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project.

Yours sincerely

Pp [redacted]

Dr Sabita Uthaya
Chair

E-mail: nrescommittee.london-riverside@nhs.net

Enclosures: *List of names and professions of members who were present at the meeting and those who submitted written comments*

"After ethical review – guidance for researchers"

Copy to: *Ms Becky Ward, Imperial College London*
Lynis Lewis, noclor. CNWL Research and Development

NRES Committee London - Riverside

Attendance at Committee meeting on 01 December 2014

Committee Members:

Name	*Profession*	*Present*	*Notes*
Ms Anna Bischler	Pharmacist	No	
Dr Ben Braithwaite	General Practitioner	Yes	
Mr John Clifford	Dean	Yes	
Dr Julian Collinson	Consultant Cardiologist	Yes	
Ms Stephanie Ellis	Former Civil Servant	Yes	
Dr Matthew Hyde	Research Scientist	No	
Dr Margaret Jones	General Practitioner	Yes	
Ms Alexandra Mancini	Modern Matron	Yes	
Ms Fanny Mitchell	Retired NHS Manager	Yes	
Mr Kamen Shoylev	Lawyer	Yes	
Mrs Dinah Smith	Retired Head Teacher	Yes	
Dr Sabita Uthaya	Consultant Neonatologist	No	
Ms Julia Williams	Senior Producer	No	
Dr Daniel Wood	Clinical Psychologist	Yes	

Also in attendance:

Name	*Position (or reason for attending)*
Miss Tina Cavaliere	REC Manager
Miss Hattie Ismail	In-house CRA (Clinical Research Associate)

Appendix B

Sociodemographic information of children's secondary caregivers

Table 49

Sociodemographic Information for Children's Secondary Caregivers

Characteristic	*n* (%), unless stated
Gender[a]	
Female	16 (7)
Male	230 (93)
Age (years), mean (SD)[a]	37.7 (6.9)
Relationship to child[a]	
Biological father	229 (93)
Biological mother	9 (3.5)
Adoptive mother	1 (0.5)
Adoptive father	1 (0.5)
Stepfather	1 (0.5)
Maternal grandmother	3 (1)
Other	2 (1)
Ethnicity[a]	
Asian	24 (98)
Black	21 (9)
Mixed	9 (4)
White	181 (74)
Other	11 (4)
Relationship status[a]	
Married/civil partnership/cohabiting	228 (93)
Divorced/widowed/legally separated	3 (1)
Single and none of the above	8 (3)
In relationship but not cohabiting	7 (3)
Employment status[a]	
Working for an employer	191 (78)
Self-employed	44 (17)
Looking after the home and family	10 (4)
Retired	1 (1)
Highest qualification[b]	
GCSE or lower	34 (14)
College (e.g., A levels, NVQ, BTEC)	50 (21)
Graduate	161 (65)

a $N = 246$
b $N = 245$ due to missing data for this variable

Appendix C

Phase 1: Participant Information Sheet and Consent Form

Imperial College London

Central and North West London NHS
NHS Foundation Trust

Healthy Start, Happy Start

1

Phase 1: Participant Information Sheet

Version 3.0 dated 31 July 2015

You are being invited to fill in a questionnaire as part of a research study looking at child behaviour and development. The study is being run by Dr Paul Ramchandani and his team at Imperial College London.

Why have I been invited to take part?

We are inviting you to take part because you have an infant/child aged **approximately 12-36 months.** We are asking lots of parents to fill in this questionnaire, as we are interested in hearing from a range of parents/caregivers about their child's behaviour.

What will I have to do?

If you are interested in taking part in the study simply complete the enclosed **consent form** and short **questionnaire** (including your contact details). You can then return them both to us via your health professional (if you received this from a health professional, such as your health visitor) or send them back to us directly in the freepost envelope provided. If you would prefer, you can complete the questionnaire and consent form online by following the link to the study website: **www.happystart.org.uk**

This questionnaire is widely used either in standard care or research studies. Based on your responses to the questionnaire, there is a chance you may be asked if you would like to take part in the next phase of the research project, which involves a programme designed to help parents/caregivers to better understand and respond to their infant/child's communication and behaviour. This is done through play and interaction sessions, where a member of the study team would meet with you and your child and explore different ways of responding to common situations.

If this happens we will contact you and give you further information on the next phase of the study, and you can think about whether or not you would like to take part. If you decide to complete the screening questionnaire this does not mean you have to take part in the next phase of the study. All your responses are really valuable to us.

Do I have to take part?

Participation is entirely voluntary, so it is up to you whether or not you would like to fill in the questionnaire. If you do decide to take part in any aspect of the study and later change your mind, you are free to withdraw at any time without giving a reason. Whether or not you take part in the study will not affect your healthcare or that of your infant/child in any way.

What will happen to my information?

All your information will be kept strictly confidential. Your information will be stored securely on Imperial College London sites and computers and will only be accessible by some members of the research team. Your responses will be kept separately to your personal details so they will not be identifiable to anyone outside of the research team.

Thank you for taking the time to read this information. If you have any questions about the study or the questionnaire please contact us, we would be delighted to talk about what the study involves.
020 8383 4167 or happystart@imperial.ac.uk

NHS National Institute for Health Research *The Healthy Start, Happy Start project is funded by the National Institute for Health Research's HTA Programme*

Imperial College London

Central and North West London NHS
NHS Foundation Trust

Healthy Start, Happy Start

2

Phase 1: Consent Form

Please read the following questions carefully, by initialing the boxes below you are giving your consent to complete the screening questionnaire and to be contacted about the second phase of the study.

Version 4.0 dated 29 October 2015

Please initial box

I confirm that I have read and understood the Information Sheet (version 3.0, dated 31/07/2015)	☐
I have had the opportunity to consider the information. If I have had any questions I have been able to contact the research team to discuss these, and they these have been answered satisfactorily.	☐
At this stage of the study I am agreeing to complete the screening questionnaire.	☐
I understand that there is a separate, second phase to the research, and I agree to being contacted about that research if I am suitable to take part.	☐
I understand that taking part is voluntary and that I am free to withdraw from the study at any time, without having to give a reason and without it affecting my or my child's medical care.	☐
I understand that sections of any of my research notes may be looked at by responsible individuals from the NHS Trust, Imperial College London or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to access my records that are relevant to this research.	☐
I understand that my personal data will be stored securely and may be used for future related research.	☐

PARENT/CAREGIVER:

.....	.....	.....
NAME	SIGNATURE	DATE

NHS National Institute for Health Research

The Healthy Start, Happy Start project is funded by the National Institute for Health Research's HTA Programme

Site ID								

Screening ID								

Appendix D

Phase 2: Participant Information Sheet and Consent Form

Imperial College London

<<Local Trust Logo to be Added>>

Healthy Start, Happy Start

Phase 2: Participant Information Sheet
Version 5.0 dated 3rd July 2017

You are being invited to take part in a research study, which aims to help parents with their children's behaviour. The study is being run by Professor Paul Ramchandani and his team at Imperial College London, alongside your local NHS trust.

What is the study about?
We are looking at a home-based programme that aims to help parents/caregivers to better understand and respond to their child's communication and behaviour. We want to find out whether this programme is helpful for families and if so, whether it's more helpful than the support that is currently available.

Why have I been invited to take part?
You are invited to take part because you have a child aged between **12-36 months**. We would like to see families from a variety of different backgrounds and day to day family lives. Parents/caregivers can be involved with their child either on their own, or with the child's other parent/caregiver. We are interested in finding out about young children's behavior. We are particularly interested in how parents and children share every day activities together and whether our programme is helpful in thinking about both the positive and challenging moments.

Do I have to take part?
Participation is voluntary, so it is up to you whether or not you would like to take part. Before you decide if you would like to take part, it is important that you know what the study is about and what taking part would involve. We would be delighted to answer any questions you have about the study.

You may wish to discuss the study and taking part with other people such as your partner, family, friends, GP or health visitor. If you do decide to take part and later change your mind, you are free to withdraw from the study at any time and without giving a reason. If you withdraw from the study, we will not collect any new information about you but will ask you if we can keep previously collected data.

What will I have to do?
If you would like to take part you will join one of two groups: one group will receive the programme (the programme group) and the other group will continue to receive their usual care (the monitoring group). You will join one of these groups at random using a specially designed computer system to ensure that there is a 50/50 chance of joining either group. This is done so that every family has the same chance of receiving the programme or continuing to receive usual care that is already available on the NHS or from other local services, keeping things fair for all families involved. Both groups are essential to the study to really understand how, and if this programme is useful for families.

What's involved for all families?
Families in both the programme group and the monitoring group will have three research visits during the study from our research team. The first visit will be once you agree to take part in the study, the second visit will be 5 months later, and the final visit will be 2 years later.

During these research visits you will complete some questionnaires, and we will have a chat about your child's behaviour. All these questionnaires are widely used in standard care or in research studies. We will also bring along some toys and film some short clips (only a few minutes long) of you doing some activities together with your child (e.g., playing together).

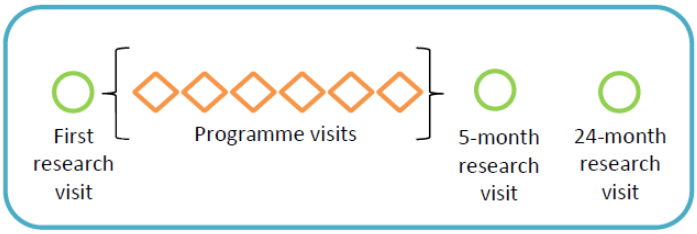
At the two year visit we will observe your child playing with some games and toys. . During this visit, we will ask for your permission to contact your child's nursery worker/teacher to ask them to fill out a questionnaire too, as we know these settings can be an important part of your child's life. We will also ask if we can take a mouth swab from the inside of your child's cheek for the genetic part of our research. This will be done by brushing a swab (like a cotton bud) on the inside of your child's cheek. DNA, which makes up our genes and can influence our growth and development, will be extracted from the swab and analysed to look at the activity and type of key genes related to children's development and behaviour (the names of some genes that have been shown to be potentially relevant include: BDNF, FKBP5, NR3C1, DRD4, SLC6A4 and 5-HTTLPR). At present this information is limited in what it can tell us about the health of individual children but its inclusion in the study will help us understand more about how genes are involved in children's development and behaviour.

We would also like to build a broader picture of how children in the study are developing by asking for your permission to add information that is routinely collected about your child's health and education from their NHS and school records. This is a common feature of large-scale research studies. It makes it easier to collect information about questions that may be difficult to remember (e.g., the number of visits your child has had with their health visitor) and may help to improve future services and policies.

For both groups, if two parents/caregivers are taking part in the study, it would be important for you both to attend all the visits together. All visits will be at your home, or at another convenient and suitable location if you prefer.

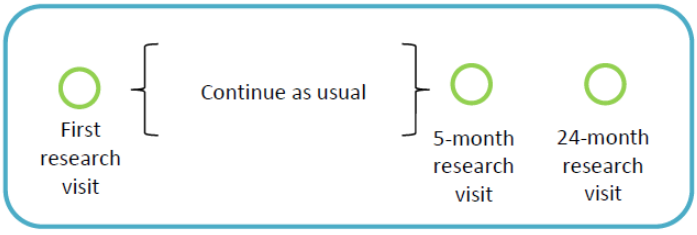
What's involved if we join the programme group?

If you join the programme group you will also receive 6 home-based programme visits from the same member of our programme team, who is a trained health professional. During these visits they will film some short clips of you and your child doing everyday activities together, such as playing with toys, reading together and **mealtimes**. They will then look back at these clips with you at your next visit, talking about what you can notice together about your child's behaviour and communication. They will also help you to try out different ways of responding to day to day situations, as well more challenging moments. These visits will happen approximately every 2 weeks at a time that suits you. During your time on the programme, you and your family can continue to receive any other medical care that you have been receiving or wish to receive.



What's involved if we join the monitoring group?

If you are allocated to the monitoring group you and your child will continue to receive the standard medical care usually available to you, and we will see you at the 3 research visits that both groups receive.



What are the possible benefits of taking part?

We cannot say for certain that taking part will be of benefit to you or your child. However, the programme has been used in many research studies previously, and parents have found it helpful. It has also been shown to help parents to learn more about their child and how to respond to their behaviour.

What are the possible disadvantages and risks of taking part?

The disadvantages of taking part are likely to be small. You would need to put aside time for the visits, for both groups the 3 research visits will take approximately 2 hours each. If you are in the programme group you will also have 6 programme visits (approximately 1 hour each). Most families enjoy both types of visits, but if for any reason you were to feel uncomfortable or wished to stop, we would of course take a break or end the session. The mouth swab we ask for is not painful to collect. The swab tends to feel like the child has a straw in their mouth. You can decide not to provide a mouth swab from your child and still complete the rest of the visit.

Will my taking part in the study be kept confidential?

All your information will be kept strictly confidential. Your information will only be accessible by certain members of the research team. We will give you a unique 'family number' so that your name and contact details are not stored with any information from the study. Likewise, the sample we take from the mouth swab will be labelled with your 'family number', and not stored with any details about you or your child. We plan to analyse these samples in a UK based lab, but there is a small chance that this sample may need to be sent out of the UK for analysis. None of your personal information would be sent with this, and the sample would be returned to the Imperial College approved tissue bank at the end of the study.

If you decide to take part we will write to your GP and your health visitor with your permission, just to let them know that you are taking part in the study. Otherwise everything you say will be treated confidentially. The only exception to this would be if something you said or wrote suggested any concerns about your own or someone else's wellbeing. In this case we would be required to contact your GP or another care provider as soon as possible, to ensure you receive the support you need. Whether or not you take part will not in any way affect the healthcare you or your child receives.

What will happen to my information?

All your information will be stored securely at Imperial College London, other approved academic institutions/universities, and NHS Trust sites and will only be accessible by authorised members of the research team. Some anonymised information may be made available to support other ethically approved research, where families' have given their permission for us to do so. In accordance with Imperial College policy, we will keep your information for 10 years after the study has finished. DNA samples will be stored frozen in a secure laboratory tissue bank approved by Imperial College London, in accordance with the Human Tissue Act, and may be used for future ethically approved research, both inside and outside of the UK. We will write to you and your child to let you know about any such future studies to provide you with further information on the nature of the study as well as contact information so that you can let us know if you or your child would prefer for the sample not to be included or to be withdrawn from the bank. All research studies like this are also checked to make sure they are being run appropriately and to a high standard. If this happens it might mean that your research notes and personal information are viewed by Imperial College staff or other research bodies, and if this occurs all information will be kept confidential.

We may use some video clips or some specific quotes from your feedback, in research or training presentations, with permission. We would never use quotes in a way that would identify you or your family. While this is not possible for video clips, they would not be presented alongside any personal details. However, you can say no to this and still take part in the study, it is completely up to you.

If you give us permission to do so, information from your child's routine health and education records may be shared with us by the NHS or Department for Education. We do this by 'linking' to these records using your child's details (e.g. name, date of birth, NHS number) through a secure transfer system, so the data providers can correctly identify

your child. Before we receive the records of relevant information all personal details are removed and replaced with a unique 'information number' like a barcode, which would make the information anonymous. We will not know who is who in the database but we can use the information number to join pieces of information together for each child. We would not share any of your study information with these government departments or agencies.

What will happen to the results of the study?

We will publish the findings of the study in scientific journals and may present them at conferences. This is always done in a way that means families cannot be identified. We will also send all families a summary of our findings at the end of the study and keep them updated on our progress. We will get in contact with you if there are developments in the study, and you can also check the study website (www.happystart.org.uk) for any additions or updates to the terms of the study.

Who has reviewed the study?

All NHS research is reviewed by an independent group, called a Research Ethics Committee, in order to protect your wellbeing, rights and dignity. This study has been reviewed and approved by NRES Committee London – Riverside Research Ethics Committee.

Who is organising and funding the research?

The research is being organised by Imperial College London. The study is being funded by the Health Technology Assessment programme of the National Institute for Health Research.

Expenses and payments

You will be given a £20 voucher each time you complete one of the 3 research visits with the research team (i.e. at the beginning, after 5 months, and after 2 years) as a thank you for your time and effort. If you choose to have the visits somewhere other than your home, we will reimburse the cost of travel.

What if there is a problem or something goes wrong?

It is unlikely that anything will go wrong, but it is important that you have this information in case it does. If you are harmed due to someone's negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way in which you have been treated during the course of this study then you should immediately inform the investigator using the contact details on this information sheet. The normal National Health Service complaints mechanisms are also available to you. If you are still not satisfied with the response, you may contact the Imperial AHSC Joint Research Compliance Office.

What happens next?

If you are happy to take part we can arrange your first home research visit at a time that suits you. At this visit we will discuss the study further with you, answering any questions you may have and check you are still happy to take part.

Further information and contact details

If you would like more information about the study or would like to discuss it with one of our researchers, please phone or email us using the contact details below.

Professor Paul Ramchandani (Principal Investigator):

[redacted]

Dr Christine O'Farrelly (Trial Manager):

[redacted]

Thank you for your time and interest in this study

NHS National Institute for Health Research

The Healthy Start, Happy Start project is funded by the National Institute for Health Research's HTA Programme

Imperial College London

<<Local Trust Logo to be Added>>



Healthy Start, Happy Start

Phase 2: Consent Form

Please read the following questions carefully, by initialing the boxes below you are giving your consent to take part in the second phase of the *Healthy Start Happy Start* study. Please check the Healthy Start Happy Start website (**happystart.org.uk**) for any additions or updates to the terms of the study.

Version 5.0 dated 3rd July 2017

	Please **initial** box:
1. I confirm that I have read and understood the Information Sheet (version 5.0 dated 03/07/17).	☐
2. I have had the opportunity to consider the information and ask any questions, which have been answered satisfactorily.	☐
3. I agree to take part in the second phase of the study with my child. This means I will be visited at least three times by the research team. I understand that I will then join one of two groups, which means I may or may not receive the home-based programme described in the Information Sheet.	☐
4. I understand that participation is voluntary and that I am free to withdraw from the study at any time, without having to give a reason and without it affecting my or my child's medical care.	☐
5. I understand that sections of any of my own and my child's research notes may be looked at by responsible individuals from the NHS Trust, Imperial College London or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to access my records that are relevant to this research.	☐
6. I agree to myself and my child being video and audio-recorded for the purposes of the study.	☐
7. I agree that quotations and sections from my video and audio recordings may be used in publications, training or presentations for research purposes and agree to this. **NB:** You can say no to this, and still participate in the research	☐
8. I agree to my General Practitioner and Health Visitor being informed of my participation in the study.	☐
9. I understand that all information I provide will be kept strictly confidential and will be stored securely at Imperial College London, other approved academic institutions/universities, and NHS Trust sites, for a minimum of 10 years.	☐
10. I understand that my personal data will be stored and may be used by the study team to contact me in the future to see if I would be interested in taking part in further Healthy Start, Happy Start-related research.	☐
11. I understand that de-identified information I provide may be shared anonymously with other researchers to support studies in the future, both in and out of the UK.	☐
12. I agree for my child's nursery teacher/school teacher to be contacted to fill in a questionnaire about my child's communication and behaviour. **NB:** You can say no to this, and still participate in the research	☐

Imperial College London

<<Local Trust Logo to be Added>>

13. I agree to the collection of a cell sample using a gentle cheek swab from my child, to the extraction of DNA from the sample, and an analysis of the activity and type of genes related to child behaviour and development (e.g., BDNF, FKBP5, NR3C1, DRD4, SLC6A4, and 5-HTTLPR). I understand that the genetic tests do not have any implications for the health of my child, and the results will not be given to me. I agree to the storage and possible further analysis of my child's DNA to assess the activity levels of other genes involved in child development and behaviour, both in the Healthy Start, Happy Start study, and future ethically approved research both in and out of the UK. Due to the nature of the analysis, it is possible that both related and unrelated genes may be looked at.
NB: You can say no to this, and still participate in the research

☐

14. I understand that my child's details such as name, gender, date of birth, address and NHS number may be used to make an accurate link to the information held and maintained by the NHS, National Health Service databases (such as the Hospital Episode Statistics and Maternity and Children's Data Set), NHS Digital and other central UK NHS bodies about my child. This may be used to help contact me or provide information about aspects of my child's health status and development, including e.g., information such as hospital admissions, prescriptions, and service use. The specific criteria of these data points may be extended in future.
NB: You can say no to this, and still participate in the research

☐

15. I understand that my child's details such as name, gender, date of birth, address and education number may be used to make an accurate link to the information held and maintained by the Department for Education on the National Pupil Database about my child. This may be used to help contact me or provide information about my child's education, such as educational attainment, attendance, and development. The specific criteria of these data points may be extended in future.
NB: You can say no to this, and still participate in the research

☐

PARENT/CAREGIVER:	.....	.....	.....
	YOUR NAME	SIGNATURE	DATE
RESEARCHER:	.....	.....	.....
	RESEARCHER'S NAME	SIGNATURE	DATE

NHS National Institute for Health Research

The Healthy Start, Happy Start project is funded by the National Institute for Health Research's HTA Programme

Appendix E

Administration manual for the doll's house play story stem battery

Version Number	Date	Changes made	Made by
1	20 October 2017	N/A	COF, BB, PR
2	28 March 2018	Changes pertain mainly to two review meetings held 14 February and 27 March 2018. These include: • Adding advice on setting up the environment (pg. 3). • Holding back on prompts/reflections questions to allow children to develop the play (pg. 3). • Use of feelings questions and special cases which may permit not including a/some feelings questions (pg. 3 and 4). • Managing situations where the child may become preoccupied with telling their own story (pg. 3). • Use of issue prompts to allow scope for these to be used where the child only briefly develops the play/or does not address the problem. • Clarification provided regarding the 'excluded' child for the feelings question in scenario 3 (pg. 9). • Guidance on when to 'bring the child back' to the problem or to progress to the next question (pg. 4).	COF, BB
3	9 May 2018	Changes pertain to a review meeting held on 8 May 2018. These include: • Guidance on verifying the story with children who do not verbalise during the interview (pg. 4). • Use of issue prompts to engage children who may struggle/be resistant to finishing the story without the interviewer's assistance (pg. 5) Use of 'wondering aloud' issue prompts for children who may not respond to direct questioning (pgs. 8, 10, 11).	BB

Background

The story stems used in the HS,HS study at the two year follow up visit were developed to elicit children's narratives and representations concerning two dimensions: *parent-child relationships* (including parental discipline, parental sensitive responding, child proximity seeking) and *peer relationships* (including prosocial responding and conflict resolution). The story stems are adapted from the MacArthur Story Stem Battery (MSSB; 2003) with some adaptations included from the Dolls House Play (Woolgar & Murray, 2010; Ramchandani, Murray et al., 2011). Their adaptation from the MSSB is in line with the developers' intention that the battery can be "tailored to the needs of individual researchers/studies" where "some researchers add stories of their own" or "mix and match" stems from other measures provided that they are carefully piloted. Piloting for the HS,HS study took place with 15 children across two nursery settings and 5 children in home settings with adjustments made iteratively until the final procedure, as described here, was reached.

Note on developmental suitability

In regards the suitability of the story stems for the HS,HS sample Bretherton and Oppenheim (2003[1]) detail that even two year olds begin to create simple stories with replicas of human figures and animals. They highlight that Wolf et al (1984) identified five increasingly complex levels of play (1) acting towards the figure or replica as if it were alive, (2) making it act toward and interact with other figures, (3) ascribing feelings and sensations to it, (4) endowing it with simple moral judgements, and finally (5) imbuing it with cognitions. By 2.6 years all children (n = 9) had mastered the first four levels and by 4 years all had attained the fifth level. Additionally, Nelson and Gruendel (1981) showed that 3 year old children correctly answered simple questions about "what happens when" concerning simple and familiar routines. Older children tended to add more detail but even 3 year olds described causally related events in the appropriate order.

Materials:

Each kit includes one of each of the following:

- 'dolls set' (see Appendix 1 Figure 1)
- 'friends set – female child (see Appendix 1 Figure 2)
- 'friends set – male child (see Appendix 1 Figure 3)
- 'breakfast set' (see Appendix 1 Figure 4)
- 'outdoor scene' (see Appendix 1 Figure 5)

Setting up the story stems

To optimise the environment for the child to engage with and develop the story stems you should try to set up the environment with the following considerations in mind:

- As part of the 'child measures' used in the HS,HS two year follow up visit try to negotiate access to a space that ensures that parent is not in immediate eye line of the child or could be a source of a distraction (e.g., in an adjacent room within ear shot with open door in keeping with safeguarding procedures).

[1] See for citations Bretherton, I. & Oppenheim D. (2003). The MacArthur Story Stem Battery: Development, Administration, Reliability, Validity, and Reflections about Meaning. In R. N. Emde, D. P. Wolf, & D. Oppenheim (Eds.), *Revealing the inner worlds of young children: The MacArthur Story Stem Battery and parent child narratives* (pp. 27–54). New York, NY: Oxford University Press.

- It may be helpful to ask the parents where the child is likely to be comfortable and most likely to engage. You can use cushions etc to ensure the child can sit comfortably especially in an adult child.
- Ideally the best location includes a table as it reduces the likelihood that the dolls will become unbalanced. The type of table is not important (e.g., you could be seated on the floor and use a coffee table).
- It is important to clear the area and immediate environs of all materials and anything that could conceivably act as a distraction regardless of how unattractive it may seem at first!

Prompts

Avoid inviting the child to participate using direct questions e.g., "Would you like to pick a doll?" "Would you like to move on to the next day?" as child may say "No". Instead, try **"Let's pick a doll."**, "**Let's move on to the next day."**

In general it is helpful to balance the use of prompts/questions against opportunities for the child to develop the play scenario. This may mean inhibiting an instinct to 'jump in' with a reflection about the play or holding back on progressing to a question or issue prompt. If you feel the situation calls for some input from the researcher you might consider using a non-descriptive vocalisation such as "ahh or hmmm" to minimise intrusion yet still contributing to the momentum of the play.

Some children intervene in the presentation of one or more story stems. If this occurs, the interviewer can tactfully say **"I'll tell the beginning and you get to finish it"**. If the child interrupts again you can use the following techniques to ask the child to **"Wait for the story"/"Can you listen out for the story?"/"So now it's [RA's name] turn", "Listen and wait".** If a child misinterprets the set-up of the activity or is intent on telling their own story instead/is otherwise distracted it may be helpful to invite the child to help you with 'your stories' **"we might be able to do [that e.g., picnic] story when we are finished, first I'd really like your help with my story, remember I tell the beginning and you get to tell me how it ends".**

Prompts during the child's response are required if a child moves the figures in ways that are ambiguous. For example, when a child moves two figures against each other and it is unclear if hitting or hugging is meant the interviewer may ask **"What are they doing?"** Similarly if the child speaks for an unidentified protagonist the interviewer can ask "**Who said that?"** If the child replies "I did" the interviewer can say **"Who in the story said that?"** If a child moves the dolls without providing any verbal explanation, for example if a child moves a doll to wipe the floor without saying anything, the interviewer can offer some narration, while also seeking verification from the child by saying **"It looks like [Mummy] is cleaning the juice up, is that right?"** Occasionally a child may ask for an additional prop used in a previous story. If this occurs say **"Just pretend."** If the child insists the interviewer can say "**You'll get to play with it after we finish the stories."** The prompt **"Anything else?"** can be used if the initial response is very brief. Each scenario has two "issue prompts". These are important as the child may avoid a scenario or create an unrelated story. The interviewer can use the issue prompts twice (if necessary) to ascertain the extent to which the child is unwilling to address the story problem. It can be tricky to determine when to interject with an issue prompt, re-set the story (e.g., in scenario 3 between the child showing you how the children play and setting up

the problem), and/or move to the next question. In these cases it may be helpful to first allow the child scope to develop the play and then when it has become clear that the story has become unrelated (e.g., fantasy theme such as space) or repetitive to use an issue prompt or progress to the next question. A caveat here is to be broad minded about what might be related/relevant for coding. [2]

Some children may find the focus on them to tell the next part of the story, without the interviewer's assistance, discouraging. Sometimes children will ask the interviewer to tell the end of the story instead and, on these occasions, the researcher can say **"I wondered if you could help me finish the story, as I'm not quite sure how it ends"**. The interviewer can also remind the child that the story telling is a collaborative effort by saying **"We can tell these stories together, I'll tell the beginning, and then you can tell me what happens next".**

Feelings questions

For each scenario there are a number of 'feelings' questions where the child is asked how the protagonist and other figures feel in relation to the scenario/problem. There tends to be wide variation in children's ability to answer these questions. As such it can be helpful to remain vigilant to children's responding in case adherence with the script (i.e. progressing through the prescribed feelings questions) appears to inhibit children's level of engagement. For example children may say 'I don't know' or not provide a response in relation to a feelings question and appear to lose confidence when asked further feelings questions. In these (small number) of cases you may consider not including a feelings question in subsequent scenarios if you believe it is affecting the child's engagement. You should still proceed if the child offers a response that does not qualify as a feeling (e.g., if when asked about how the child feels about the spilled juice the child says 'yummy' or if the child simply says 'good' or 'bad' as these can still be coded for valence (i.e. positive of negative responding)). Thus it is necessary to exercise judgement in this aspect of the script erring on the side of following the script as intended (i.e. asking all questions) wherever possible.

Interviewer reactions

Avoid saying 'good job', instead you can say **'thank you'**. Similarly avoid saying "that was a great story" because it is difficult to use these remarks after disjointed or aggressive stories. Instead you can say **"I can see you're working really hard on these stories, thank you".** It is important to remain non-judgemental when negative or chaotic responses to stories are given.

Repeating the child's utterances

It is often helpful to repeat what children, especially young children, say during the narratives. Doing so not only conveys to them that they are heard but also helps transcribers understand narratives spoken in a quiet or unclear voice. Most young children accept these repetitions, however they may be unnecessary with older or very articulate children who may find the technique awkward or find it to be an interruption.

[2] Reflective meetings can be helpful here to ensure that relevant themes are not being overlooked.

Ending a story/Transitioning to a new story

Many children (especially older children) will end stories by saying “all done” or “the end”. If the child does not do so the interviewer must judge when it is time to move on. Several criteria may be used to facilitate this decision. For example if the enactment becomes repetitious the interviewer can ask **“How did it all end up/Show me what happens in the end/Is this the end?”** To lead into the next story the interviewer can say **“Now it is time for a new day. On the new day…”** For figures that are not involved in the new story you can say “the sister isn’t in the next story” and place the figure in the box where the props are kept out of sight. If CHILD seems like they’re going to struggle with transition to new scenario: **“Let’s pretend everyone’s sleeping until the next day”**. Act out figures sleeping and then wake them up for new day.

Wind down

After all the stories have been presented the interviewer invites the child to have the dolls to engage in a fun activity. At this point the child is told that they can play with any figure/prop they want. The purpose of this is to provide a pleasant and relaxed ending for the session during which no other specific demands are made of the child.

Warm up

The aim of the warm-up is for the child to:

1. Talk to the RA.
2. Pick up the dolls and manipulate them.
3. Speak for the dolls.
4. Say something that relates to the scenario at hand.

RA should introduce the activity to the child without taking the dolls out, nor any of the props. This is to allow the child to focus on identifying genuine family members (rather than the most exciting looking doll in the tin!).

> **I: So next up, we're going to look in the purple bag. I've got some dolls for you to meet in the purple bag, and we're going to use them to tell three stories. I'll tell the beginning, and you can tell the end. For these stories we're going to pretend that we're here in [CHILD'S name's] house. So you live here. Who else lives here?**

If certain of family set-up, RA can make suggestions (e.g. Does Mummy/Daddy live here?) although avoid this if any uncertainty. Unless CHILD is really insistent, try to encourage the CHILD to only pick people who live in their house (rather than extended family members, friends, or imaginary figures). If child does not select any adult caregivers prompt "Is [main caregiver] going to be in the stories?" Once the family have named their family members RA can run through a summary of their family.

e.g. I: Okay, so we have [CHILD], Mummy, Daddy, and [sibling name(s)]. Shall we pick dolls for everyone now? [Once dolls are picked]

> **I: Okay so we have three stories to tell but before we do that I think the family could have a party. I wonder what kind of party they could have.**

If the CHILD responds develop the CHILD'S play until the CHILD has fulfilled 1-4 above. If the CHILD is in need of help tell a story. You can wonder aloud if helpful. If the CHILD wants to sing "happy birthday" by all means join in and sing along.

For example, if no response:

> **I: "I wonder if it could be an ice cream party."**
>
> **If CHILD has still not picked up dolls or engaged, RA pick up dolls**
> **I: "[CHILD] has some chocolate cake, I wonder who she/he gives the cake to."**
>
> **If CHILD picks a doll**
> **I: "What does she/he say about the lovely cake?"**

If the CHILD is manipulating and talking for the dolls immediately, the party can be short-lived and the spilled juice scenario can be moved to relatively quickly. Silly play and over-expression is encouraged during the warm-up to allow the CHILD to feel comfortable enough to engage in play behaviours. For the warm-up, it is okay to use demonstrations for the CHILD or issue leading prompts, however these should not be used for the subsequent story stems.

1. Spilled Juice




RA should start this scenario with the materials for the Breakfast Set (kitchen table, chairs, and jug) out of sight.

> **I: "Okay, now the party has finished and it's a new day. We're going to pretend the new day is in your house. And the family are going to have breakfast."**
>
> **I: "We're in [CHILD'S NAME]'s house. Here's the table, and here's the chairs."** **Get kitchen table and chairs out**. **Let's put the chairs around the table. Let's sit the family down."** **If possible, encourage CHILD to help you sit them down and choose seats**
>
> **I: "So who do we have around the table?"** **After CHILD responds** **"That's right, we have [CHILD'S NAME], Mummy, Daddy [or MAIN CAREGIVERS], and [as appropriate, e.g. sister doll, Jack, girl doll, Grandma etc.]."** *If CHILD switches roles at this point (e.g. switches sister doll and CHILD doll around, or CHILD doll to Mummy doll, reconfirm with CHILD until it's established which doll is which).*
>
> **I: "The family are thirsty and they are going to have some juice."** **Place juice jug on table*.* **"Here's the family drinking their juice. Now, [CHILD'S NAME] gets up and reaches across the table and Uhoh! She/he spilled her/his juice all over the floor.** **Make child spill the pitcher onto the floor so that it is visible to the CHILD. Put lots of emotion into voice.**
>
> **I: "Show me and tell me what happens now."**
>
> *[Allow some time for the child to show you what happens before asking the next questions]*
>
> **I: "How does [CHILD'S NAME] feel?" Why does [CHILD'S NAME] feel [emotion, e.g. sad, happy]."**
>
> **I: "How does [MAIN CAREGIVER] feel?" "Why does [MAIN CAREGIVER] feel [emotion]?"**
>
> **I: "Show me what happens in the end. How does it all end up?"**

Prompts

*If CHILD doesn't respond/moves the dolls on to something unrelated/only very briefly develops the play particularly if they have not addressed the problem use the **issue prompt***	**What happens about [CHILD'S NAME] spilling the juice?**
If CHILD seems discouraged by direct questioning, but has not developed the play, the interviewer can 'wonder aloud' the issue prompt	**Hmm, I wonder what happens about [CHILD] spilling the juice.**
*If CHILD continues to move dolls onto something unrelated (e.g. "and then they all go back up to bed") try to bring the child back once more with a **second issue prompt** (only if necessary) prompt*	**[CHILD'S NAME] spilt the juice. What happens about [CHILD'S NAME] spilling the juice?**

after first allowing the child to develop the scenario to ensure it is unrelated	
If CHILD only picks up props	**Did anything else happen?**
If CHILD picks up the jug themselves without speaking for the dolls	**The jug is back on the table. How did that happen?**
If CHILD says "Not sure" or "I don't know".	**That's okay**

2. Hurt Knee



a) Child 1 and Child 2 playing



b) Child 1 tripping

I: "Okay so now we're going to move on to a new day. On this new day, your family are outside so we can put the kitchen away" **Place kitchen in bag and any additional family members, beyond main caregiver(s), and study CHILD, back in with the mix of other dolls*.*

I: On the new day we are going to pretend that [CHILD'S NAME] is outside in the park. "So we've got
Mummy and Daddy [or MAIN CAREGIVERS] over here. And then we have [CHILD'S NAME] here." **Place a little distance away from main caregiver*.*

I: "Now we need to pick a friend for [CHILD'S NAME]. Who's your best friend?" **Wait for CHILD to respond** **"Okay, let's pick a [FRIEND NAME] doll."** This seems to be a point where CHILDREN get a little distracted, or reserved. Try to encourage them to pick one doll from the Family Set. Once the child has picked a doll move the box so that it is out of sight.

If CHILD picks a sibling/character from previous family scenario **"The sister isn't in the next story" "We need to pick a friend doll for [CHILD'S NAME]** instead**?"** If CHILD insists that their sibling is the friend doll that is fine.

If CHILD doesn't pick a doll for their best friend, RA can pick friend doll and this should be the same gender as the CHILD and be selected from the Family Set. If CHILD picks the friend doll and this is of a different gender then that's fine. If CHILD picks doll but struggles to name their best friend then use standard names:

Child 1. *CHILD'S NAME.*
Child 2. *Jack/Emily/CHILD'S choice.*

Once friend has been picked and assigned a name

I: "Okay, so we have [CHILD'S NAME] and [Jack/Emily/CHILD'S CHOICE] and they're playing outside. Suddenly, [CHILD'S NAME] runs, and trips and falls and cuts their knee. "OW! I've *really* hurt my knee!" **Show the child doll running, tripping and falling as speaking. Put lots of emotion into voice.**

I: "Show me and tell me what happens now."

I: "How does [CHILD'S NAME] feel? Why does [CHILD'S NAME] feel [emotion, e.g. sad, happy]."

I: "How does [MAIN CAREGIVER] feel? Why does [MAIN CAREGIVER] feel [emotion]?"

I: "Show me what happens in the end. How does it all end up?"

Prompts

If CHILD doesn't respond/moves the dolls on to something unrelated/only very briefly develops the play particularly if they have not addressed the problem use the ***issue prompt***	**What happens about [CHILD'S NAME] hurt knee?**
If CHILD continues to move dolls onto something unrelated (e.g. "and then they all go back up to bed") it's okay to let the child develop this part of play briefly. Then use a ***second issue prompt***	**[CHILD'S NAME] "ow my knee is really sore" What happens about [CHILD'S] hurt knee?**
If CHILD seems discouraged by direct questioning, but has not developed the play, the interviewer can 'wonder aloud' the issue prompt	**Hmm, I wonder what happens about [CHILD'S] hurt knee.**
If CHILD begins to develop play	**Did anything else happen?**
If CHILD picks up the doll themselves without speaking for the dolls	**[CHILD'S NAME] is standing up. How did that happen?**
If CHILD is moving dolls but not speaking, or saying little, RA can confirm with CHILD what is happening e.g. If [CHILD 2] goes towards [CHILD 1] on ground and picks other child up.*	**It looks like [Jack/Emily/CHILD'S CHOICE] has helped [CHILD'S NAME] up, is that right?**
If CHILD says "Not sure" or "I don't know".	**That's okay**

3. Three's a Crowd



a) Child 1 and Child 2 playing



b) Child 3 asking to play

I: "Okay so now we're going to move on to a new day."

I: I: "On this day [CHILD'S NAME] doll is in the park again. Mummy/Daddy/[MAIN CAREGIVER(S) is/are over here." *Caregivers to be a slight distance from child dolls.*

I: "[CHILD'S NAME] is playing with a friend called [Harry/Ava], and they're playing here with [Harry/Ava's] new ball. Show me how they play with the ball." *Allow CHILD* a few seconds of playing with the ball. **"[CHILD'S NAME] and [Harry/Ava] are playing with the ball."**

I: *RA to move CHILD 3 whilst talking.* **"[Charlie/Sophie] comes over and says "Can I play with you?""**

I: *RA to pick up CHILD 1 whilst talking.* **"[CHILD'S NAME] says "Sure!""**

I: *RA to pick up CHILD 2 whilst talking.* **"[Harry/Ava] says "No way! I don't want to play with [Charlie/Sophie]."**

I: "Show me and tells me what happens now."

I: "How does [CHILD'S NAME] feel?" Why does [CHILD'S NAME] feel [emotion]?

I: " How does [excluded child[3]] feel? Why does [NAME] feel [emotion]?

I: "Show me what happens in the end. How does it all end up?"*Prompts*

If CHILD doesn't respond/moves the dolls on to something unrelated/only very briefly develops the play particularly if they have not addressed the problem use an ***issue prompt***	**What happens about [Charlie/Sophie] wanting to play with the ball?**
If CHILD doesn't respond, or moves the dolls on to something unrelated use a ***second issue prompt.***	**But I said I don't want [Charlie/Sophie] to play with us. What happens about who gets to play?**
If CHILD seems discouraged by direct questioning, but has not developed the play, the interviewer can 'wonder aloud' the issue prompt	**Hmm, I wonder what happens about [Charlie/Sophie] getting to play.**
If CHILD begins to develop play	**And then what happens?**
If CHILD moves the dolls without speaking for the dolls	**What's happening now?**
e.g. Or CHILD brings all three children together and kicks ball between them.	**It looks like the children are playing with the ball together, is that right?**
If CHILD is resistant to more direct questioning.	**I wonder how [CHILD'S NAME] feels**
If CHILD says "Not sure" or "I don't know".	**That's okay**

[3] Note. The subject can change to include any character depending on who becomes excluded in the scenario. If CHILD develops the scenario such that no child is excluded ask how Charlie/Sophie feels (i.e. the child originally targeted for exclusion).

Appendix 1: Materials



Figure 1. Dolls set



Figure 2. Friends set – female child



Figure 3. Friends set – male child

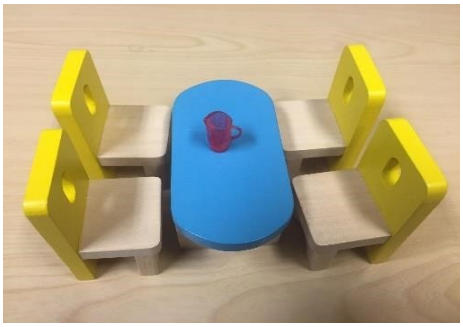


Figure 4. Breakfast set



Figure 5. Outdoor scene

Appendix F

Coding manual for the doll's house play story stem battery

BEHAVIOUR PROBLEMS – *Actions from/descriptions of <u>any</u> dolls, directed towards any character, prop or object*	
Aggressive behaviours	
Score	**Description**
0	**No aggressive behaviours**
1	**Minimal aggressiveness** (*rough play gone too far, verbal threats, push out of the way*)
2	**Moderate aggressiveness** (physical aggression shown but only transiently and not intensely; *kicking, hitting*)
3	**Severe aggressiveness** (*character attacks another, hitting over and over; fighting, aggression resulting in injury or death*). Account for force when scoring aggression (e.g. one hit that sends doll flying off table may be a '3')
Ostracising, bullying, or excluding behaviours	
Score	**Description**
0	**No ostracising behaviours**
1	**Minimal ostracism** (inconsiderate behaviours that lead to the exclusion of a character (*e.g. Harry and TC move away from Charlie to play football with no acknowledgement of him*) but more passive than higher scores (i.e. no explicit incidence of 'you can't play').
2	**Moderate ostracism** (teasing or making fun of character, or demeaning comments. Moderate ostracism could include a verbal exclusion (*'you can't play'* or *'she doesn't get to sit at the table'*) but less persistent or intentional than a '3'.
3	**Extreme ostracism** (overt verbal or physical exclusion, *"No way, you can't play, go away"*, or kicking/pushing doll away from play. The exclusion is reinforced and it is made clear that the intention is to reject character.
Destructive behaviours	
Score	**Description**
0	**No destructive behaviours**
1	**Minimal destructive behaviours** (*character is careless; spills juice again* (even repeatedly); *knocks over chair*)
2	**Moderate destructive behaviours** (*character pops ball, character breaks jug on floor,* character's actions cause permanent damage)
3	**Extreme destructive behaviours** (*character is intently and consciously causing damage, e.g. character smashes up kitchen)*
Prosocial behaviours	
Score	**Description**
0	**No prosocial behaviours**
1	**Minimal prosocial behaviours** (characters friendly towards each other, but not going out of their way to include or help; e.g. *I clean up the juice; TC and Charlie play ball* (but nothing about how this was resolved)
2	**Moderate prosocial behaviours** (characters explicitly helpful, considerate, comforting or friendly, behaviour is clearly benefitting others; e.g. *friend takes child over to parents when hurt knee; comforts when child is upset; helping Mum to clean up juice, everyone has a turn with the ball or shares the juice, Dad gets a plaster*)
3	**Extreme prosocial behaviours** (characters go above and beyond to help; characters modify their behaviour to help (e.g. *Daddy drank all the juice, so I go to the shop to get some more; Ava isn't being nice so child tells her she needs to apologise to Sophie, and then they all play together; Mum rescues child off the cliff*)

Behaviour	**SCORING**			
	SS1	SS2	SS3	**TOTAL**
Aggressive behaviours [AggBP]				
Ostracising behaviours [OstBP]				
Destructive behaviours [DesBP]				
Prosocial behaviours [PSoBP]				

NARRATIVE COHERENCE	
Score	**Description**
0	**No response** or "I don't know what happens". Child may repeat a portion or all of the story stem without any additions.
1	**Not coherent.** There are fragmented shifts in story line. Child does not return to original story stem or address the conflict.
2	**Partially coherent.** Child exhibits an understanding of the conflict but does not offer any resolution when a resolution is expected or does not offer an ending to the story. Alternatively, the child may handle the conflict by changing the constraints presented in the original stem (e.g. the juice never spilt, they didn't hurt their knee). A portion of the narrative may be incoherent.
3	**Moderately coherent.** Child exhibits an understanding of the conflict but only offers a partial resolution, narratives may appear fragmented or include incoherent shifts. Child may shift the focus of the story away from original dilemma (e.g. sibling spills juice repeatedly instead) or indirectly by offering an easier solution (e.g. child walks away without ball)
4	**Coherent.** Child demonstrates an understanding of the conflict or story and offers a resolution with some story embellishment; a segment of the story may be incoherent.
5	**Very coherent.** A very coherent, logical, sequential series of events that are related to the story stem. Child may add to the story but does not change the original story stem. An understanding of the conflict and a resolution to the conflict are presented. There are no incoherent shifts in the story and there is a lot of embellishment.

SCORING					
	SS1 *[NCSS1]*	**SS2** *[NCSS2]*	**SS3** *[NCSS3]*	**TOTAL** *[NCTOTAL]*	

SELF-REGULATION	
[CONCSr] **Sustains concentration; willing to try repetitive tasks, pays attention during instructions** (*stays on task)*	
0	Child able to concentrate and persist with task, even toward end of stories
1	Child occasionally distracted but generally persistent, does not require prompt from assessor
2	Child's attention frequently drifts, and examiner provides multiple prompts
3	Child not able to concentrate or persist on much of the task *(spends most of the time off-task)*

[SIGSr] **Distracted by sights and sounds** *(objects/other toys around them, people or noises, dolls/table)*	
0	Child does not become distracted by sounds and sights in room
1	Child occasionally distracted but refocuses attention with minimal *(none or only one)* prompt
2	Child occasionally distracted but refocuses attention and returns to task with several prompts
3	Child is frequently distracted by the environment, and has trouble focusing even with assessor help

[DESSr] **Is careless or destructive with test materials** *(including dolls, table and chairs, ball, jug, and bags)*	
0	Child is not careless and is not destructive
1	Child is careless or slightly destructive one time (can include dropping object on floor "by mistake")
2	Child is repeatedly careless but does not damage materials (*e.g. flings table or dolls*)
3	Child gets "too rowdy" with material *or is intentionally forceful with the materials* (*e.g. throws doll across room; high force or potential for damage*); or breaks or damages test materials.

[TOUSr] **Refrains from indiscriminately touching materials** *(grabbing at bags, dolls out of researcher's hands)*	
0	Child shows self-restraint even with interesting materials and does not prolong packing up/moving on
1	Child generally shows self-control but reaches for attractive objects once or twice during assessment
2	Child repeatedly grabs at test materials, and/or needs multiple reminders from researcher not to touch objects, *or to stop grabbing*
3	Assessment often interrupted by child's difficulty with grabbing, touching materials

[INTSr] **Lets examiner finish before starting task; does not interrupt**	
0	Child waits before pointing to materials, reaching for *dolls*, *telling story*; is not impulsive
1	A few *[one or two]* instances of impulsive behaviour (*e.g. interrupts researcher delivering SS, insists "we need a Grandad doll"*).
2	Child is often impulsive across multiple stems or highly impulsive during one story stem
3	Child impulsive throughout assessment, needing lots of boundary-setting

[SEASr] **Remains in seat appropriately during task**	
0	Does not get out of seat or climb up and down (occasional adjustment in body position is appropriate; or to pick something up that has fallen)
1	Gets out of seat once (including sliding off chair), returns to seat if prompted
2	Leaves seat multiple times, but listens and responds to prompts if given
3	Out of seat frequently or difficult to manage (e.g. runs around room, climbs on furniture), *requires frequent prompts, and/or does not listen to assessor about returning to seat*

Appendix G

Intraclass correlation coefficients for story stem coding

Table 50

Intraclass Correlation Coefficients for Story Stem Coding

		95% confidence interval	
	Intraclass correlation	Lower bound	Upper bound
Thematic content			
Aggression	.98	.97	.99
Destructive	.91	.86	.95
Bullying/exclusion	.98	.95	.98
Prosocial	.92	.87	.96
Narrative coherence			
Coder A & B	.95	.89	.97
Coder A & C	.81	.64	.90
Coder B & C	.68	.42	.83
Behavioural regulation			
Coder A & B	.91	.87	.95
Coder A & C	.85	.77	.90
Coder B & C	.75	.62	.83

Appendix H

Participant sociodemographic information and variable descriptives disaggregated by chapter subsamples

Table 51

Sociodemographic Characteristics of Samples by Inclusion in Chapter

Characteristic	Full sample – Ch.2 (N = 300)	Subsample – Ch.3 (n = 261)	Subsample – Ch.4 (n = 256)	Subsample – Ch.5 (n = 185)
Child characteristic				
Sex (female), *n* (%)	137 (46)	122 (47)	120 (47)	84 (45)
Age at baseline (months), mean (SD)	23.0 (6.7)	23.0 (6.7)	23.2 (6.7)	23.2 (6.8)
Age at 5MFU (months), mean (SD)	29.2 (7.1)	.	.	29.1 (7.4)
Age at 2YFU (months), mean (SD)	47.8 (7.3)	48.0 (7.1)	47.8 (7.1)	48.4 (7.5)
Ethnicity, *n* (%)				
Asian	17 (6)	15 (6)	14 (5)	11 (6)
Black	18 (6)	14 (5)	15 (6)	15 (8)
Mixed ethnicity	61 (20)	55 (21)	54 (21)	37 (20)
White	194 (65)	170 (65)	166 (65)	114 (62)
Other	10 (3)	7 (3)	7 (3)	8 (4)
Primary caregiver characteristic				
Sex (female), *n* (%)	287 (96)	250 (96)	245 (96)	175 (95)
Age at baseline (years), mean (SD)	34.2 (5.8)	34.7 (5.5)	34.7 (5.4)	34.0 (5.8)
Parental status, *n* (%)				
Biological or adoptive mother	288 (96)	251 (96)	246 (96)	176 (95)
Biological or adoptive father	12 (4)	10 (4)	10 (4)	9 (5)
Ethnicity, *n* (%)				
Asian	31 (10)	27 (10)	26 (10)	19 (10)
Black	18 (6)	16 (6)	16 (6)	15 (8)
Mixed ethnicity	22 (7)	18 (7)	18 (7)	16 (9)

Characteristic	Full sample – Ch.2 (N = 300)	Subsample – Ch.3 (n = 261)	Subsample – Ch.4 (n = 256)	Subsample – Ch.5 (n = 185)
White	217 (73)	191 (73)	187 (73)	125 (68)
Other	12 (4)	9 (4)	9 (4)	10 (5)
Relationship status, *n* (%)				
Married/civil partnership/cohabiting	255 (85)	227 (87)	222 (87)	150 (81)
Single	29 (10)	21 (8)	21 (8)	21 (11)
In relationship but not cohabiting	11 (3)	8 (3)	8 (3)	9 (5)
Divorced/widowed/legally separated	5 (2)	5 (2)	5 (2)	5 (3)
Employment status, *n* (%)				
Employed	162 (54)	143 (55)	140 (55)	93 (50)
Looking after home and children	112 (37)	93 (36)	91 (36)	75 (41)
Paid parental leave	16 (5)	16 (6)	16 (6)	10 (5)
Full-time student	10 (4)	9 (3)	9 (3)	7 (4)
Highest qualification, *n* (%)				
Graduate	191 (64)	176 (67)	172 (67)	112 (60)
A-level/NVQ/BTEC	78 (26)	62 (24)	61 (24)	49 (27)
GCSE or lower	31 (10)	23 (9)	23 (9)	24 (13)

Table 52

Descriptive Statistics of Measures of Child Behaviour, Parental Psychopathology and Parental Discipline of Samples by Inclusion in Chapter

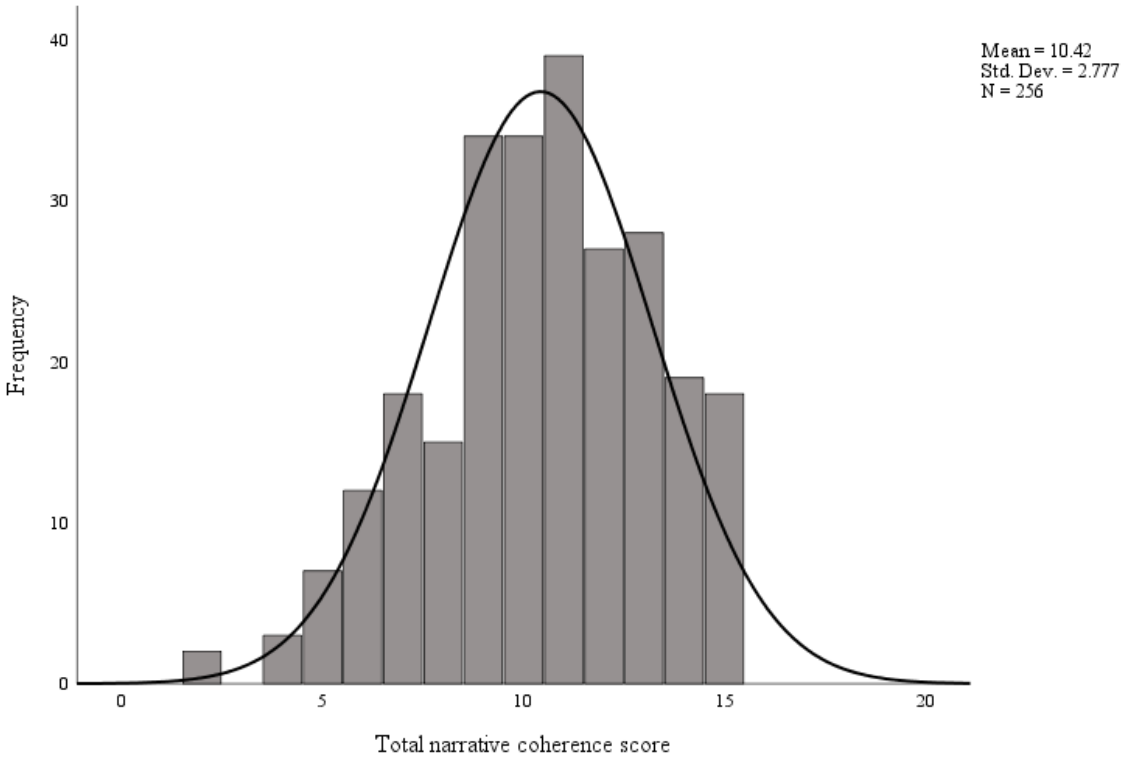
Characteristic, *M* (*SD*)	Full sample – Ch.2 (*N* = 300)	Subsample – Ch.3 (*n* = 261)	Subsample – Ch.4 (*n* = 256)	Subsample – Ch.5 (*n* = 185)
Child characteristic				
Behaviour - CBCL	41.7 (21.4)	40.2 (20.6)	40.5 (20.6)	42.2 (20.6)
Behaviour - PPACS	32.9 (9.9)	32.4 (9.4)	32.6 (9.4)	31.9 (10.5)
Behaviour – SDQ	13.9 (4.7)	13.6 (4.6)	13.7 (4.6)	14.1 (4.6)
Parenting characteristic				
Parental depression – PHQ-9	4.3 (4.2)	3.9 (3.7)	4.0 (3.7)	4.1 (4.2)
Parental anxiety – GAD-7	4.8 (4.3)	4.7 (4.3)	4.8 (4.3)	4.5 (4.1)
Parent lax discipline	3.7 (0.4)	3.7 (0.4)	3.7 (0.4)	3.7 (0.4)
Parent overreactive discipline	3.9 (0.5)	3.9 (0.5)	3.9 (0.5)	3.9 (0.5)

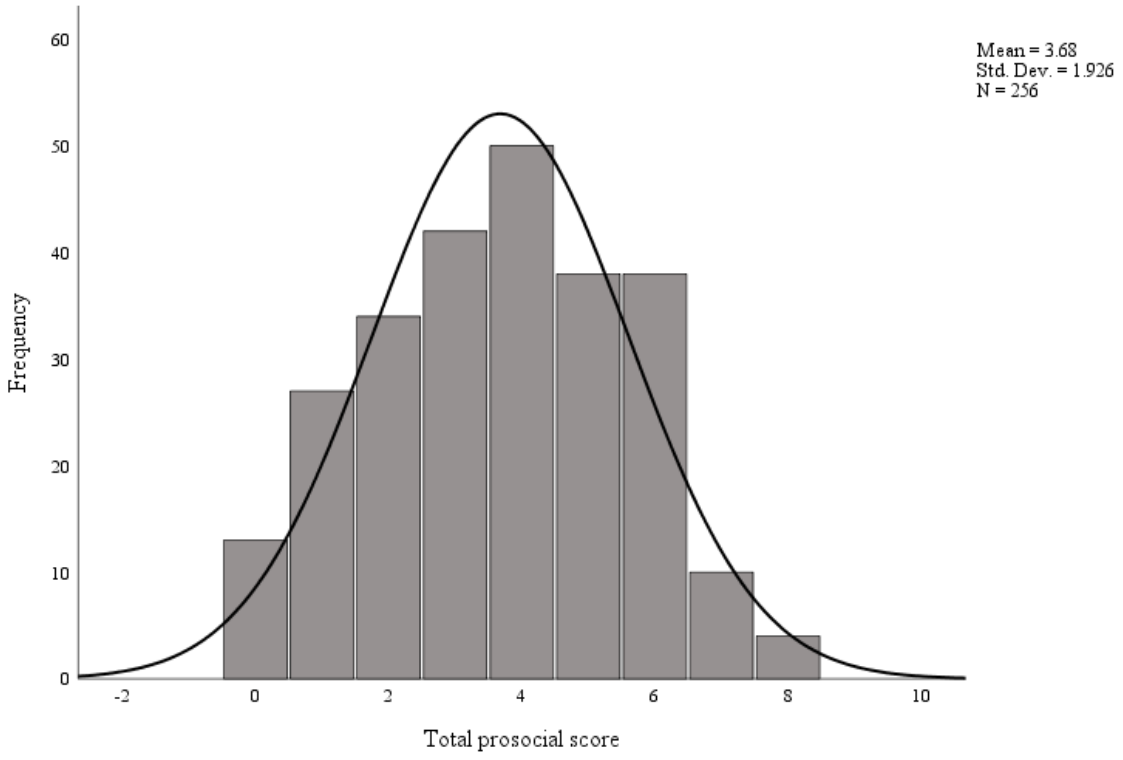
CBCL = Child Behavior Checklist; PPACS = Preschool Parental Account of Childhood Symptoms; SDQ = Strengths and Difficulties Questionnaire; PHQ-9 = Patient Health Questionnaire-9; GAD-7 = Generalised Anxiety Disorder-7; Parent lax discipline = Parenting Scale subscale; Parent overreactive discipline = Parenting Scale subscale

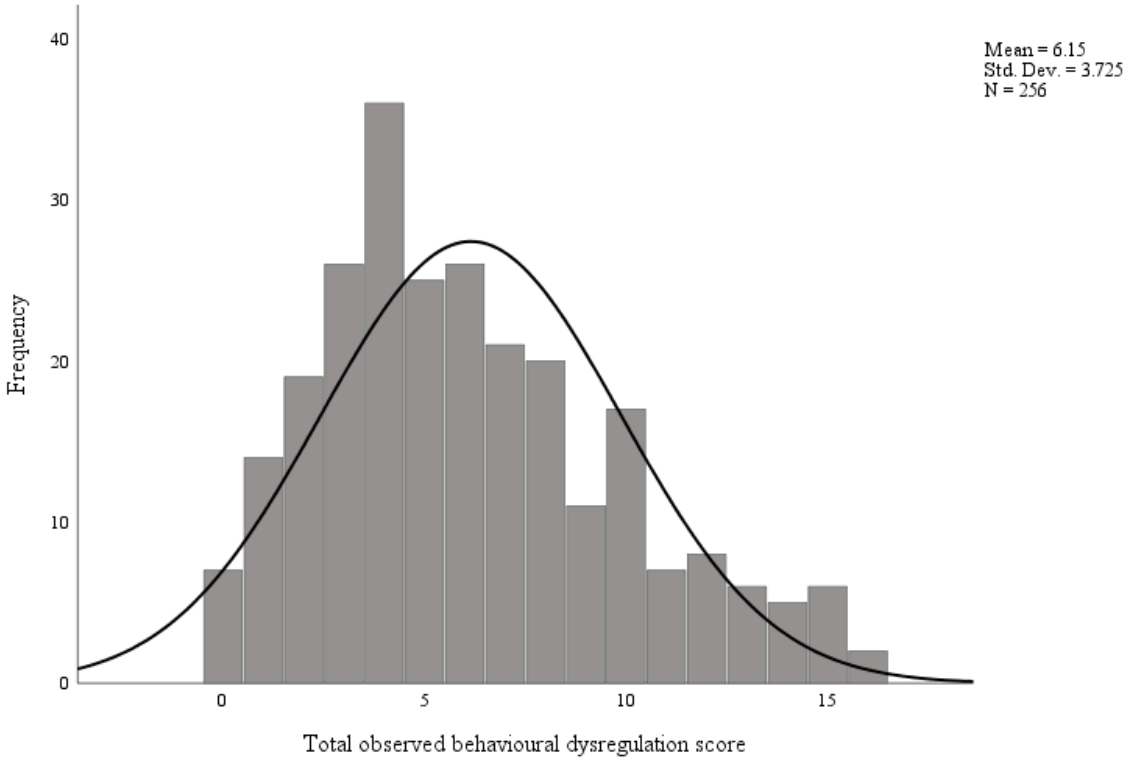
Appendix I

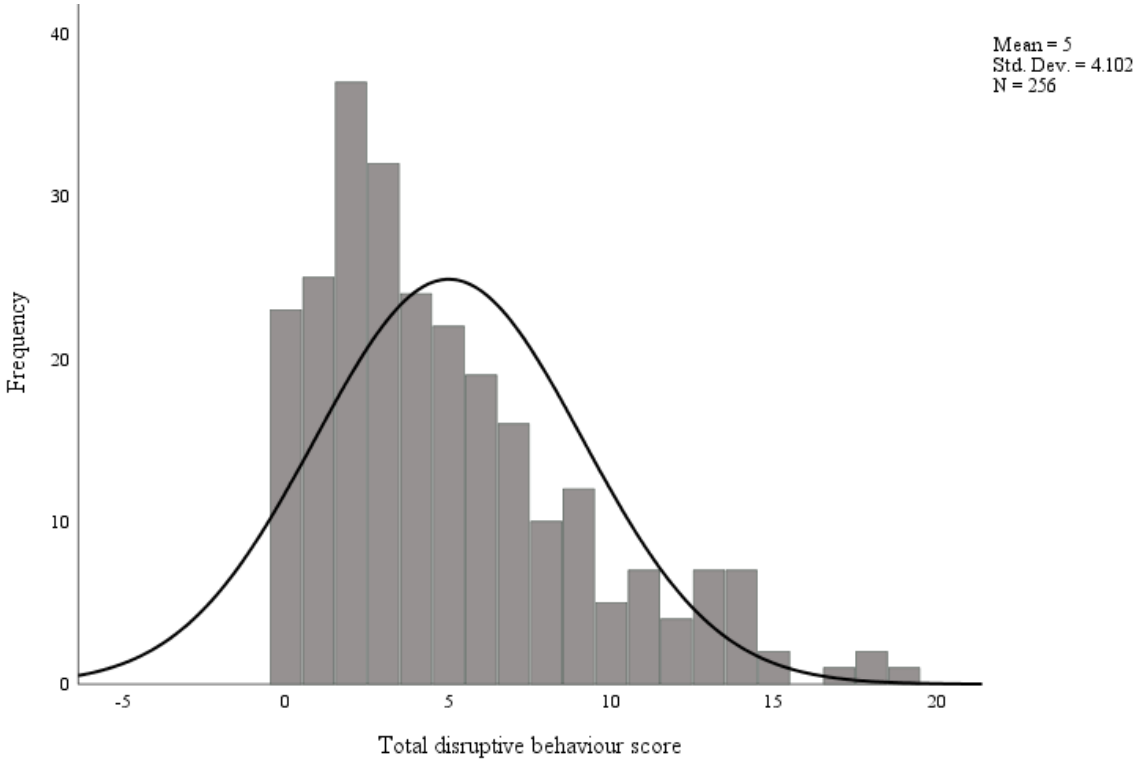
Data testing univariate assumptions of normality

Histograms illustrating the distributions of constructs from children's story stem data.






Mean = 6.15
Std. Dev. = 3.725
N = 256
Frequency
40
30
20
10
0
0
5
10
15
Total observed behavioural dysregulation score


Mean = 5
Std. Dev. = 4.102
N = 256
Frequency
40
30
20
10
0
-5
0
5
10
15
20
Total disruptive behaviour score

Appendix J

Bivariate correlations between teacher-reported child behaviour and parent-related variables

Table 53

Bivariate Correlations Between Teacher-Rated Child Behaviour and Parental Mood, Anxiety, Discipline Styles, and Educational Attainment

Variable	1	2	3	4	5	6	7	8
1. Teacher-rated SDQ – total	-							
2. Teacher-rated SDQ - externalising	.86**	-						
3. Teacher-rated SDQ - prosocial	-.60**	-.57**	-					
4. Parental depressive symptoms	.11	.10	-.07	-				
5. Parental anxiety symptoms	.07	.09	-.04	.78**	-			
6. Parental overreactive discipline	-.12	-.08	.13	.07	-.00	-		
7. Parental lax discipline	-.14	-.16*	.00	-.05	-.04	-.02	-	
8. Parental educational attainment	-.19**	-.16*	.13	-.13*	-.15*	-.15*	.06	-

Note. * $p < .05$, ** $p < .01$

SDQ = Strengths and Difficulties Questionnaire; parental depressive symptoms = Patient Health Questionnaire-9 (PHQ-9); parental anxiety symptoms = Generalised Anxiety Disorder-7 (GAD-7; parental overreactive and lax discipline = Parenting Scale.

Appendix K

Distributions of story stem constructs split by child gender

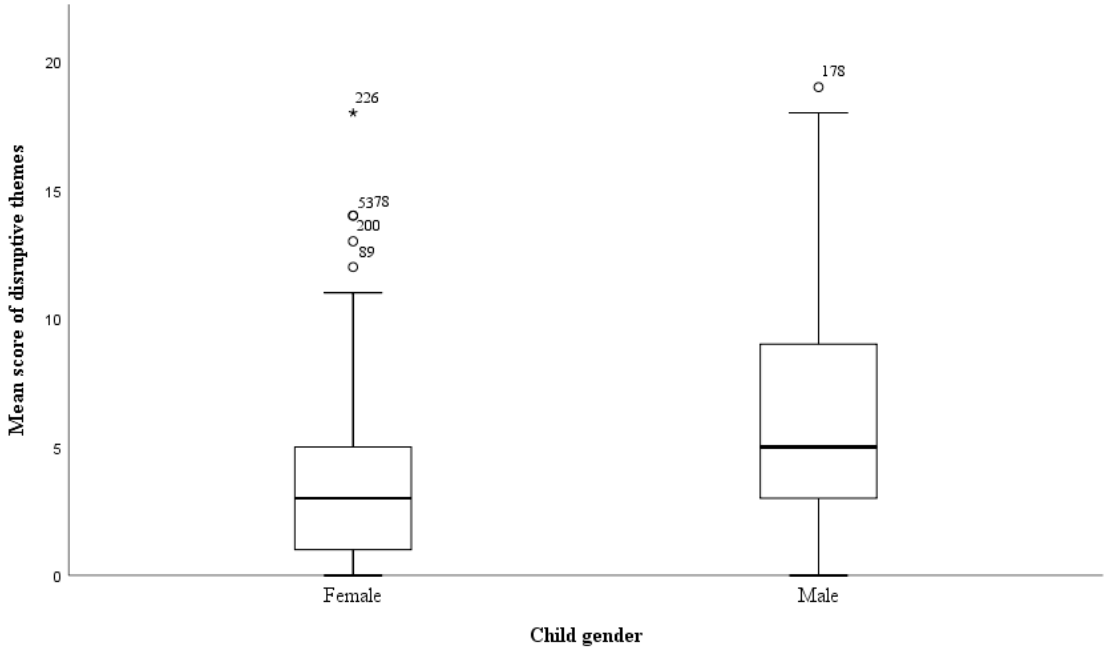


Figure 20. Distribution of mean disruptive theme scores split by child gender.

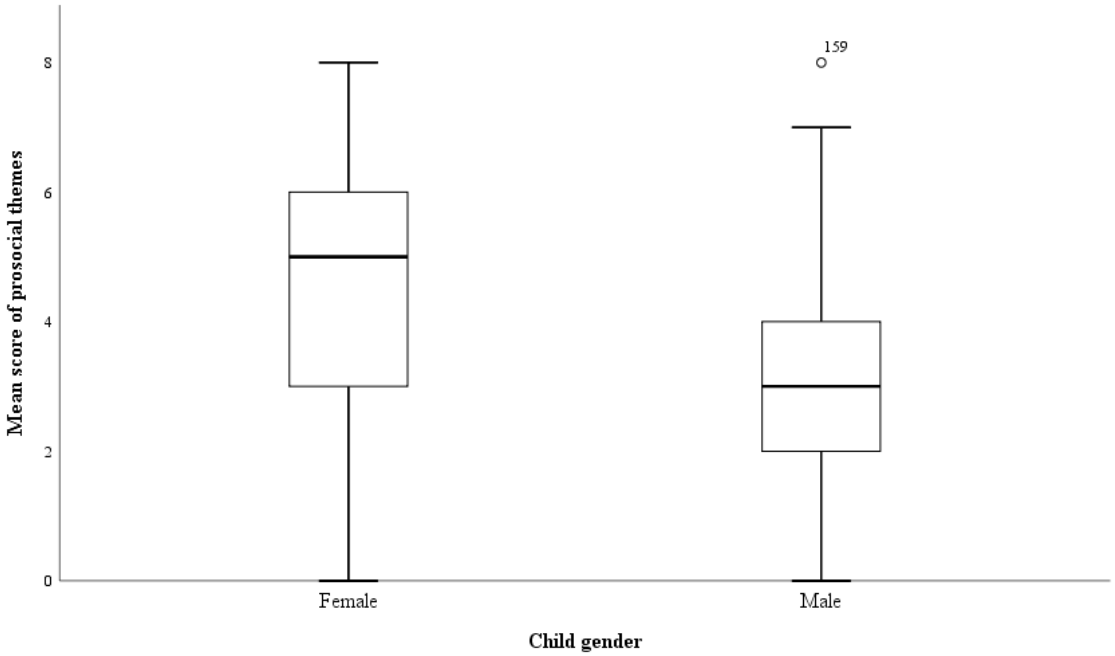


Figure 21. Distribution of mean prosociality scores split by child gender.

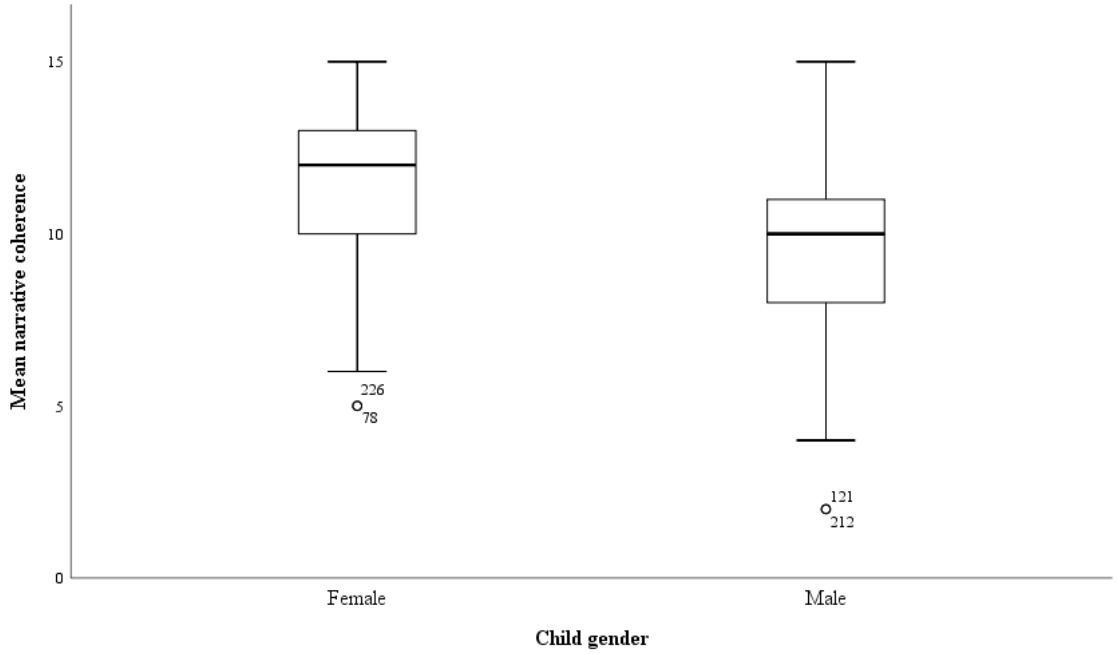


Figure 22. Distribution of mean narrative coherence scores split by child gender.

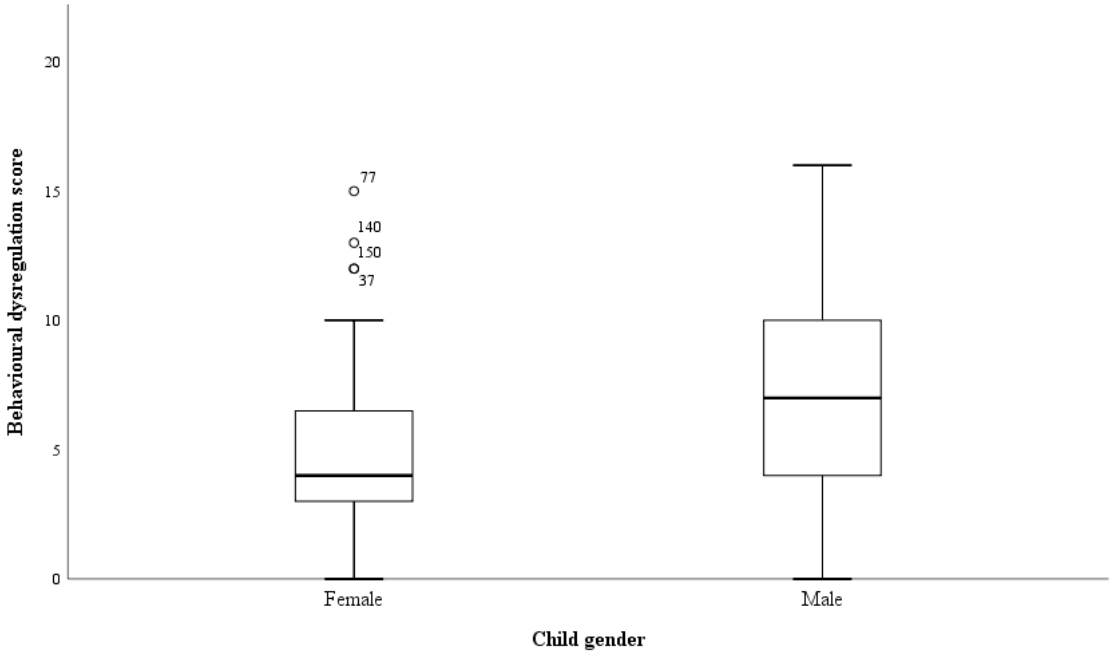


Figure 23. Distribution of behavioural dysregulation scores split by gender.

Appendix L

Distributions of story stem constructs by child age

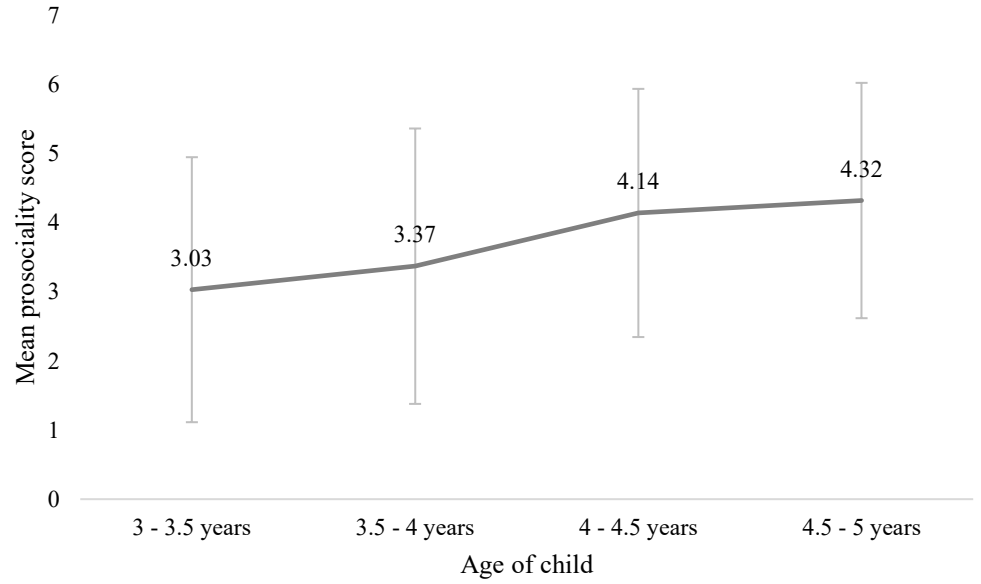


Figure 24. Means scores of prosociality by age of child.

Note. Error bars represent the standard deviation (SD).

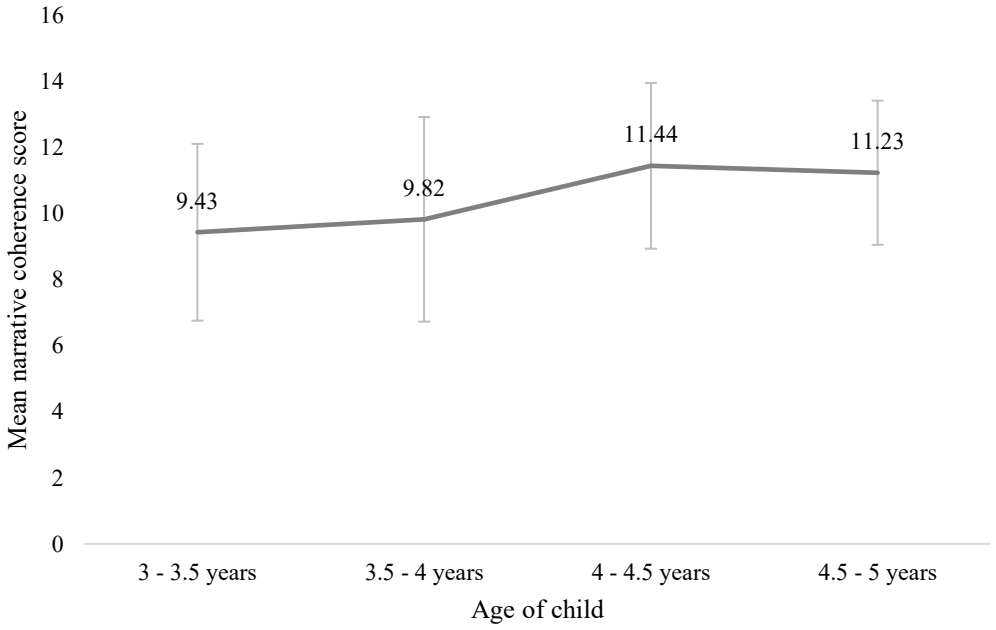


Figure 25. Means scores of narrative coherence by age of child.

Note. Error bars represent the standard deviation (SD).

Appendix M

Secondary analysis: Early child behaviour as a predictor of children's disruptive themes

Table 54

Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Expression of Disruptive Themes Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	2.22 (-2.62, 7.07)	2.46		0.90	.37
	Child gender	2.53 (1.57, 3.49)	0.49	.31	5.18	.000
	Child age	0.03 (-0.04, 0.10)	0.04	.06	0.97	.34
	Caregiver education	-0.82 (-1.58, -0.07)	0.38	-.13	-2.15	.03
	Caregiver mood	0.07 (-0.06, 0.21)	0.07	.06	1.05	.30
	Trial arm	-0.07 (-1.03, 0.89)	0.49	-.01	-0.15	.88
2	(Constant)	5.59 (-2.57, 13.75)	4.14		1.35	.18
	Child gender	2.47 (1.50, 3.43)	0.49	.30	5.03	.000
	Child age	0.04 (-0.03, 0.11)	0.04	.07	1.04	.30
	Caregiver education	-0.84 (-1.60, -0.07)	0.39	-.13	-2.16	.03
	Caregiver mood	0.07 (-0.07, 0.20)	0.07	.06	1.00	.32
	Trial arm	-0.07 (-1.03, 0.89)	0.49	-.01	-0.14	.89
	Overreactive parenting	-0.20 (-1.29, 0.88)	0.55	-.02	-0.37	.71
	Lax parenting	-0.71 (-1.82, 0.41)	0.57	-.08	-1.24	.22
3	(Constant)	3.93 (-4.40, 12.27)	4.23		0.93	.35
	Child gender	2.39 (1.43, 3.36)	0.49	.29	4.88	.000
	Child age	0.04 (-0.03, 0.11)	0.04	.07	1.10	.27
	Caregiver education	-0.75 (-1.51, 0.02)	0.39	-.12	-1.92	.057
	Caregiver mood	0.04 (-0.10, 0.18)	0.07	.03	0.53	.60
	Trial arm	-0.12 (-1.08, 0.84)	0.49	-.01	-0.24	.81
	Overreactive parenting	-0.25 (-1.33, 0.83)	0.55	-.03	-0.45	.65
	Lax parenting	-0.67 (-1.79, 0.44)	0.57	-.07	-1.19	.24
	Child behaviour	0.05 (-0.01, 0.10)	0.03	.11	1.75	.08

Note. R^2 = .13 for Step 1, ΔR^2 = .01 for Step 2 (*p*s = .44), ΔR^2 = .01 for Step 3 (*p*s = .08)
Caregiver mood = Patient Health Questionnaire-9 (PHQ-9) total score at baseline
Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline
Child behaviour = Preschool Parental Account of Childhood Symptoms (PPACS) total score at baseline

Appendix N

Secondary analysis: Early child behaviour as a predictor of children's observed dysregulation

Table 55

Linear Model of Early Parental Discipline and Behaviour Problems as Predictors of Children's Observed Behavioural Dysregulation Two Years Later

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	6.40 (2.08, 10.71)	2.19		2.92	.004
	Child gender	2.75 (1.89, 3.61)	0.44	.37	6.33	.000
	Child age	-0.05 (-0.11, 0.02)	0.03	-.09	-1.47	.14
	Caregiver education	-0.58 (-1.26, 0.09)	0.34	-.10	-1.71	.09
	Caregiver mood	0.01 (-0.11, 0.13)	0.06	.01	0.23	.82
	Trial arm	-0.39 (-1.25, 0.46)	0.44	-.05	-0.91	.37
2	(Constant)	4.61 (-2.68, 11.90)	3.70		1.25	.21
	Child gender	2.75 (1.89, 3.61)	0.44	.37	6.27	.000
	Child age	-0.04 (-0.11, 0.02)	0.03	-.08	-1.32	.19
	Caregiver education	-0.055 (-1.23, 0.13)	0.35	-.10	-1.59	.11
	Caregiver mood	0.02 (-0.10, 0.14)	0.06	.02	0.28	.77
	Trial arm	-0.38 (-1.24, 0.48)	0.44	-.05	-0.87	.39
	Overreactive parenting	0.32 (-0.65, 1.29)	0.49	.04	0.66	.51
	Lax parenting	0.06 (-0.94, 1.06)	0.51	.01	0.12	.91
3	(Constant)	4.78 (-2.71, 12.28)	3.81		1.26	.21
	Child gender	2.76 (1.89, 3.63)	0.44	.37	6.25	.000
	Child age	-0.04 (-0.11, 0.02)	0.03	-.08	-1.32	.19
	Caregiver education	-0.56 (-1.25, 0.13)	0.35	-.10	-1.60	.11
	Caregiver mood	0.02 (-0.11, 0.15)	0.06	.02	0.33	.75
	Trial arm	-0.38 (-1.24, 0.49)	0.44	-.05	-0.86	.39
	Overreactive parenting	0.33 (-0.64, 1.30)	0.49	.04	0.66	.51
	Lax parenting	0.06 (-0.95, 1.06)	0.51	.01	0.11	.91
	Child behaviour	-0.01 (-0.05, 0.04)	0.03	-.01	-0.20	.84

Note. R^2 = 16 for Step 1, ΔR^2 = .002 for Step 2 (*p*s = .80), ΔR^2 = .00 for Step 3 (*p*s = .84)
Caregiver mood = Patient Health Questionnaire (PHQ9) total score at baseline
Overreactive and lax parenting = subscale scores of the Parenting Scale at baseline
Child behaviour = Preschool Parental Account of Childhood Symptoms (PPACS) total score at baseline

Appendix O

Post hoc analysis: Group allocation as a predictor of children's behaviour at the 5-month follow-up for full HSHS sample

Table 56

Linear Model of Predictors of Post-Intervention (Five-Month Follow-Up) Child Behaviour (CBCL) for Full HSHS Sample

	Variable	*B* (95% CI)	SE B	β	*t*	*p*
1	(Constant)	16.88 (5.62, 28.15)	5.72		2.95	.003
	Child age	0.62 (0.26, 0.98)	0.18	.20	3.42	.001
	Child gender	2.45 (-2.35, 7.25)	2.44	.06	1.01	.32
2	(Constant)	-0.27 (-7.51, 6.96)	3.68		-0.07	.94
	Child age	-0.01 (-0.24, 0.22)	0.12	-.00	-0.11	.92
	Child gender	2.09 (-0.91, 5.10)	1.53	.05	1.37	.17
	Baseline behaviour	0.78 (0.71, 0.85)	0.04	.79	20.95	.000
3	(Constant)	1.43 (-6.14, 8.99)	3.84		0.37	.71
	Child age	-0.02 (-0.25, 0.21)	0.12	-.01	-0.15	.88
	Child gender	1.88 (-1.13, 4.89)	1.53	.05	1.23	.22
	Baseline behaviour	0.78 (0.70, 0.85)	0.04	.79	20.87	.000
	Group allocation	-2.26 (-5.27, 0.75)	1.53	-.05	-1.48	.14

Note. R^2 = .043 for Step 1, ΔR^2 = .583 for Step 2 (*p*s = .000), ΔR^2 = .003 for Step 3 (*p*s = .140)

Baseline behaviour = Child Behavior Checklist (CBCL) total

N = 285