The Development of the "P-Factor" in Early Childhood: Examination of Genetic and Environmental Indicators.

Melissa Barry

D.Clin.Psy. thesis (Volume 1), 2018

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Name: Melissa Barry

Date: 9/07/18

Overview

This thesis is focused on furthering the current understanding of psychopathology in the very early stages of life. The first part of the thesis aims to address the gap in literature surrounding the efficacy of interventions for preschool anxiety. The first systematic review and meta-analysis of the literature was conducted. A range of interventions were identified and considered alongside potential bias identified from quality assessment. The meta-analysis supported the use of psychological interventions as treatment for preschool anxiety, although there was significant heterogeneity in studies. Limitations as a result of the number of included studies are discussed, as well as recommendations for future research.

The empirical paper, which forms the second part of the thesis, moves away from traditional diagnostic approaches which form the basis for previous intervention research, to utilise a model that reconceptualises psychopathology into a general psychopathology dimension ("p-factor") and additional specific dimensions. This study is the first to utilise an adoptive cohort study to examine the development of the "p-factor" over time, and from the earliest age to date. Specifically it examines the strength consistency of the p-factor over development and the contributions of genetic and environmental risk indicators at each time point.

Finally, the third part of the thesis contains a critical appraisal of the research process. It provides a reflective account of the relevance of psychologists in addressing issues relating to the conceptualisation of psychopathology and the significance of findings from the field of genetics to clinical psychologists.

Impact Statement

The findings from this research have relevant clinical and academic implications for the field of developmental psychopathology. Part one of the thesis offers the first attempt to systematically analyse the efficacy of interventions targeting anxiety disorders in preschool children. The findings from this review support the use of interventions for this age group and contribute to the evidence base in support of the provision of early interventions in clinical practice. Additionally, promising effects were found for the provision of preventive interventions to preschool children showing signs of anxiety or high levels of behavioural inhibition, which have wider implications for public health policy and the education sector.

For academia, this review highlighted the lack of attention paid to preschool anxiety interventions on research agendas outside of Australia, where all of the largest randomised controlled trials were conducted. Given the demonstrated efficacy of interventions for this age group, it is important that research in this area is prioritised by other countries to ensure ecological validity in other health care systems.

The empirical paper provided an important contribution to the understanding of the p-factor, which has many academic and clinical implications. Firstly, it provided evidence of the replicability of the p-factor at the youngest age to date. Early identification of the presence of the p-factor in very young children is important, as it raises the possibility of the p-factor being utilised to identify and target children that may be at the most risk of later psychopathology.

Secondly, the p-factor model offers a dimensional approach to the conceptualisation of psychopathology which overcomes some of the many criticisms of traditional psychiatric nosology systems. In light of these findings, research may benefit from a shift in focus to more trans-diagnostic mechanisms and mechanism-

focused interventions for psychopathology. Future research employing an individual symptom level exploration of the p-factor would complement the present findings, elucidating specific casual pathways of the p-factor and different symptom trajectories over development.

Additionally, this was the first p-factor study to examine the influence of genetics on the development of the p-factor. This present study showed some, albeit modest and inconsistent, evidence of genetic influences on the p-factor in young children. Whilst possible that these differences may reflect random statistical variability, it could be interpreted as evidence of genetic innovation during toddlerhood, which could imply a high risk period for children with high genetic risk. It would be important for future research to replicate and advance these findings through examination of gene-environment interactions.

Finally, this study found evidence of significant environmental contributions to the development of child psychopathology, notably the impact of adoptive maternal psychopathology on the child p-factor. This research follows others in highlighting the important impact of maternal depression and anxiety on child psychopathology. This has clinical implications for mental health professionals, and social workers, as it serves to emphasise the importance of early recognition of maternal mental health difficulties and early access to treatment, especially to prevent detrimental outcomes for the child.

Table of contents

Overview 3

Impact Statement 4

Acknowledgements 10

Part 1: Literature review

Interventions for preschool child anxiety: a systematic review and meta-analysis.

Abstract 12

Introduction 13

Method 18

Results 20

Discussion 38

References 46

Part 2: Empirical paper

The Development of the "P-Factor" in Early Childhood: Examination of Genetic and Environmental Indicators.

Abstract 54

Introduction 56

Method 68

Results 78

Discussion 86

References 97

Part 3: Critical appraisal

Introduction 113

Conclusion 119

References 120

Appendices

Appendix A 123

Appendix B 124

Appendix C 125

Appendix D 130

Appendix E 132

Appendix F 133

Appendix G 134

Appendix H 139
Appendix I 140
Appendix J 141
Appendix K 145
Appendix L 149

List of tables and figures

Part 1: Literature review

Figure 1. PRISMA flowchart 1 123

Figure 2: Measurement decision tree 124

Table 1: Outcome measures and key findings for included studies 125

Table 2: Study characteristics 24

Figure 3.Forest plot of included studies 29

Figure 4: Subgroup forest plot of intervention vs. prevention 30

Figure 5. Subgroup forest plot by client group 31

Figure 6. Subgroup forest plot of CBT vs. other 32

Figure 7. Summary of quality assessment tool 33

Figure 8. Trim and fill funnel plot for included studies 37

Figure 9. Trim and fill funnel plot for intervention studies 38

Part 2: Empirical paper

Table 1. Demographic information. 69

Figure 1. Preschool CBCL bifactor model 130

Figure 2. School age CBCL bifactor model. 131

Figure 3. Birth mother bifactor model. 132

Figure 4.Diagram of CFA and CBM models. 77

Table 2. Summary of CFA and CBM goodness of fit statics. 79

Figure 5. ECV of child p-factor over development. 80

Table 5. Summary of ECV for CFA and CBM 133

Table 6. CBCL item factor loadings at 18 months. 134

Table 7. CBCL item factor loadings at 27 months. 135

Table 8. CBCL item factor loadings at 4.5 years. 136

Table 9. CBCL item factor loadings at 7 years. 137

Table 10. Birth mother item factor loadings, 139

Table 11.Summary of genetic regressions for child p-factor. 82

Table 12. Summary of environment regressions for child p-factor. 84

Table 13. Summary of evocative gene-environment correlations. 140

Table 14. Genetic regressions for the internalising specific factor. 141

Table 15. Genetic regressions for the externalising specific factor. 143

Table 16. Environmental regressions for the internalising specific factor. 145

Table 17. Environmental regressions for the externalising specific factor. 147
Figure 6. Graph illustrating Fisher's z test 149

Acknowledgements

First and foremost, I would like to thank my supervisor Pasco Fearon for all of his support and guidance, I'm particularly grateful for his patient tutorials on behaviour genetics and Mplus. I would also like to thank Praveetha Patalay for her helpful conversations about statistical modelling and my personal tutor, Chris Baker, for his support and encouragement during the doctoral training

This thesis would not exist without the generosity of the team at the Early Growth and Development Study and the participants that gave their time to the study.

I would like to thank all my family and friends for whom I am grateful for their unwavering support throughout this doctorate. I cannot name them all but would like to give particular thanks to my dear friends Holly and Sarah for their endless encouragement, advice and humour when it was needed most. I'm grateful to my housemates, Liz Harding and Lisa Masters, for their friendship that has made training more enjoyable. I'd also like to thank my mother and grand-mother, Tracy and Jean, for their constant love and reassurance.

Finally, I would like to dedicate this to my grandad, Malcolm Pearce (1945-2009).

Part 1: Literature review.

Interventions for preschool child anxiety: a systematic review and meta-analysis.

Abstract

BACKGROUND: Preschool anxiety is linked with later life negative outcomes including psychopathology. Childhood anxiety impacts on various areas of functioning such as educational attainment and the development of peer relationships. There is a need for a systematic review examining the range and efficacy of interventions for anxiety in this age group.

METHOD: A systematic review and meta-analysis of all randomised controlled-trials of interventions targeting preschool anxiety was conducted. PsychINFO and Medline were searched on August 2017. The Cochrane Quality Assessment tool was used to assess the risk of bias of included studies.

RESULTS: Thirteen studies were included in the review, resulting in eighteen interventions included in the meta-analyses. A range of interventions delivered in different formats were identified. The meta-analysis showed a large effect in support of the efficacy of preschool anxiety interventions in comparison to controls. Subgroup analyses showed a small effect size in support of prevention studies targeting preschool children at risk of anxiety and moderate effect size of CBT treatments for children with anxiety. Quality assessment highlighted some areas of potential bias in included studies.

CONCLUSION: Whilst the heterogeneity of the included studies warrants caution, this review supports the use of psychological interventions as treatments for preschool anxiety.

Introduction

Prevalence

Whilst there is growing recognition of the importance of paying attention to preschool emotional wellbeing (National Institute for Health and Care Excellence, 2012), and of recognising early signs of such problems in children in routine clinical practice, there have been few attempts to systematically synthesise the literature regarding effective interventions for this age group. Anxiety disorders are among the most common disorders in early childhood (Whalen, Sylvester, & Luby, 2017). The literature reports varying prevalence rates. A recent cross-sectional community study of 917 parents of preschool children (aged 2–5 years) reported prevalence rates of 19.4% for anxiety disorder using a clinician administered diagnostic measure (Franz et al., 2013). Gadow, Sprafkin, and Nolan, (2001) used parent and teacher reported measures found that anxiety disorders (including Separation Anxiety Disorder (SAD) , Generalised Anxiety Disorder (GAD) and Social Phobia (SP)) were common in both community samples (3-11%) and clinic samples (0-20%) for preschool children aged between 3-6 years. Egger & Angold (2006) provide comparative prevalence data that suggest similar rates of anxiety disorders in preschool children to adolescents, with 9.4% of pre-school children meeting diagnostic criteria for anxiety disorders.

Whilst exact prevalence rates vary and seem to depend on geographic location and measurement method, there is clear evidence that anxiety disorders are present, and are relatively common, in preschool children, with the most common diagnoses being SAD, SP, GAD and specific phobia (Whalen et al., 2017).

Chronicity and impact

Prospective-longitudinal studies have found associations between diagnosed childhood anxiety disorders and psychiatric problems in later adolescence (Bittner et al., 2007; Sterba, Prinstein, & Cox, 2007). These disorders are reported not only to have an impact on the child's later outcomes but also to

impact on family functioning. One study highlighted that, relative to children without disorder, families of children with anxiety were 3.5 times more likely to report a negative impact on family functioning, with the impact from specific disorders such as GAD and SAD being comparable to impact from disorders such as ADHD and other “disruptive” disorders (Towe-Goodman, Franz, Copeland, Angold, & Egger, 2014).

Even when using more objective measures research has found a significant difference between pre-schoolers with and without anxiety when assessing impact using a global measure of children's functioning (Bufferd, Dougherty, Carlson, & Klein, 2011). Some research has also found evidence of enduring effects of preschool anxiety on children’s relationships with their peers (Danzig et al., 2013). This prospective cohort study interviewed parents using a semi-structured diagnostic interview when their child was three years old and again at six years, finding anxiety disorder at age three was the only unique diagnostic predictor of peer functioning.

Risk factors

The prevalence, chronicity and detrimental outcomes associated with preschool anxiety highlight the importance of research aimed at understanding the origins of preschool anxiety to ensure effective early interventions. Research attempting to identify early determinants of anxiety has shown preschool anxiety appears to be both multifaceted and dynamic (Hirshfeld-Becker, Micco, Simoes, & Henin, 2008; McLeod, Wood, & Weisz, 2007; Muris, van Brakel, Arntz, & Schouten, 2011; Wichstrøm, Belsky, & Berg-Nielsen, 2013) with child temperament and parenting being the main areas of focus for literature in this field.

Theoretical models tend to postulate a central role for parenting in the development of childhood anxiety, although empirically parenting styles seem to account for only modest proportions of variance (McLeod et al., 2007). In a meta-

analytic review of the literature only 4% of the variance in childhood anxiety was accounted for by parenting, with parental control being associated with a larger proportion of the variance in childhood anxiety compared to other dimensions of parenting styles, such as parental rejection (McLeod et al., 2007). Whilst the strength of the association varied greatly depending on the method of measurement of parenting style (observational measures showing a stronger association than self-report measures or interviews), it appears that parenting styles may play only a modest direct role in childhood anxiety.

More recently prospective longitudinal studies examining parental anxiety, amongst other risk factors, have offered more insight into the impact of parenting on the development of child anxiety (Muris et al., 2011; Wichstrøm et al., 2013). There is increasing recognition of two-way interactional processes in the development of anxiety, with evidence suggesting that child characteristics may affect risk for anxiety and impact on parenting. Behaviour Inhibition (BI) is one of the most widely studied temperament characteristics that researchers have considered as a risk factor for childhood anxiety (Hirshfeld-Becker, Micco, Henin, et al., 2008; Hirshfeld-Becker, Micco, Simoes, et al., 2008). Prospective studies have shown that high scores for BI are predictive of later anxiety disorders (Wichstrøm et al., 2013), with parental anxiety predicting later anxiety disorders only for those with high levels of BI.

Research need

Despite the relatively high prevalence of childhood anxiety disorders, and the likely negative trajectory of untreated anxiety (Sterba et al., 2007), there is a large gap between those needing evidence-based treatment and those receiving it in this age group (Chavira, Stein, Bailey, & Stein, 2004).

There have been previous meta-analytic reviews of RCTs of treatments for childhood anxiety, however these have not specifically focused on the preschool

period. A recent meta-analysis of 41 studies involving children and adolescents from the ages of four to 19 years found that CBT was more effective than waiting list controls, but no more effective than non-CBT treatments or treatment as usual control groups in reducing anxiety diagnosis (James et al., 2015). There were also no differences in outcome for individual, group or family/parent formats, and no statistically significant treatment gains in the remission of anxiety diagnosis at follow-up. It is important to note that this review does not exclude studies where anxiety is not the primary disorder in focus (e.g. Austism Spectrum Disorder) and uses a large range of ages, with only four studies containing children under the age of six years old.

The James et al. review also synthesises studies implementing specific treatment protocols for diagnosis-focused treatments and transdiagnostic CBT treatment protocols. In a meta-analysis examining only transdiagnostic CBT interventions for treating childhood anxiety, children in the CBT intervention were 9.15 times more likely to no longer meet the criteria for an anxiety disorder than controls (Ewing, Monsen, Thompson, & Cartwright-hatton, 2015). However there were only two studies that included children in the preschool age range in this analysis. Whilst one meta-analysis suggested that the effectiveness of CBT interventions is not moderated by the age of the child (Bennett et al., 2013), it did not include children under the ages of six years.

Whilst the effectiveness of interventions for the preschool age group has not yet been analysed in a meta-analytic review, a narrative review concluded that there were promising treatments for preschool anxiety from some small scale RCTs and controlled trials of developmentally adapted CBT and Parent–Child Interaction Therapy (Luby, 2013). It also highlighted promising outcomes from research on preventative interventions targeting BI and parental anxiety (Rapee, Kennedy, Ingram, Edwards, & Sweeney, 2010) and called for psychotherapy as a first line

treatment for anxiety in preschool children. Luby (2013) recommended that whilst the primary caregiver should be involved in all treatments, the preschool child should be the main target of the intervention.

In light of this, it is clearly important to understand the range and efficacy of interventions for anxiety in this age group. In a recent review of reviews regarding RCTs for childhood anxiety Bennett and colleagues (2015) recommended that future reviews should include program-specific comparisons, including CBT vs non-CBT approaches, which will be sought to be applied in the current review.

Thus the following research questions were developed in relation to preschool anxiety:

What are the range of interventions for treating childhood anxiety?

How effective are interventions for treating and preventing anxiety disorders?

Are these interventions more effective when they involve parents?

Are interventions using CBT techniques more effective than other interventions?

Participants: Preschool children with a mean age of less than 7 years old. Preschool was chosen to define ages before formal primary school education which was interpreted to cover up to the age of 6 years (as is true for most western English-speaking societies). To limit intervention heterogeneity and maintain focus on common anxiety problems in preschool age children, PTSD treatment studies were excluded.

Intervention: Randomised Control Trials (RCTs) that are aimed at treating or preventing childhood anxiety using psychological/behavioural interventions. There was no restriction on type of psychotherapeutic approach.

Control: Waiting list control, treatment as usual.

Outcome: Reduction in anxiety symptom measures.

Method

A systematic search was conducted of the databases PsychInfo and Medline using the following search terms: preschool*, infant*, child*, toddler*, paediatric or pediatric, anxiety, sleep* problems, sleep* disorder*, phobia, worry, behavi* inhibition, shyness, obsessive compulsive, random*, RCT, clinical trial, controlled trial. These initial search terms (limited to title and abstract only) produced over 1,300 hits in each database. Thus the terms parent*, paternal, maternal, famil*, caregiver*were added to increase search specificity. These search terms were limited by age (birth to 12 years) using the database limit function.

This left the final total of papers from the initial search 874. Relevant systematic reviews and meta-analyses identified from the initial search and reference lists from included studies were checked for relevant papers. This led to

the inclusion of 5 more papers to the overall total screened (see Figure 1, *Appendix A*).

Papers were excluded if their main intervention was not a randomised controlled trial containing a psychological intervention aimed at treating or preventing childhood anxiety disorder or symptoms. Whilst preschool aged children are the main focus of the intervention we aimed to be inclusive and included studies with children with a mean age of less than 7 years. Physical and neurodevelopmental disabilities were excluded from the review. Foreign language papers and non-peer-reviewed papers were excluded.

Data for meta-analysis was extracted from all papers. In most studies there were multiple measures of anxiety included (see Table 1, *Appendix C*). A hierarchical inclusion rule was used in these cases, whereby anxiety symptom measures were prioritised over anxiety diagnoses and behaviour inhibition measures (see Figure 2, *Appendix B*). The software Comprehensive Meta-Analysis (CMA) was used to calculate the Cohen's d effect size and standard error. In studies where there was more than one intervention group, the sample size of the control group was divided in half so that the total sample size was not inflated in the calculation of effect size standard errors. The software package STATA 14 was used to conduct the meta-analyses and tests of publication bias. Random effects models (Dersimonian & Laird, 1986) were used for all analyses to account for differences above that of random sampling error, allowing for clinical and methodological heterogeneity (Lipsey & Wilson, 2001).

As highlighted in Table 1 (*Appendix C*), each study varied both in terms of the length of time for follow-up and the outcome measures employed (specificity of anxiety construct being measured and the levels of validation the measure had received). It is worthy of note that there were some discrepancies between effect sizes reported in papers and that which is reported in the meta-analysis below. This

is because the data extracted for the current meta-analysis used post-intervention data between groups only and not pre-post comparisons, as this allows for more consistent comparisons across studies. For example, Cartwright-Hatton et al. (2011) found an effect size of d=1.01, however in current meta-analysis this was reduced to d=0.77.

The Cochrane Collaboration Quality Assessment Tool was used for assessment of the quality of the research literature across seven different domains (The Cochrane Collaboration, 2011). For the purposes of this review alterations were made in the interpretation of the criteria for blinding of participant and personnel to adapt it to the literature being reviewed. It is commonly recognised that for certain allied health professional studies some Cochrane criteria are not directly applicable (Katrak, Bialocerkowski, Massy-Westropp, Kumar, & Grimmer, 2004). Due to the nature of psychological interventions it is not possible for participants to be blind to treatment, and thus this category was rated as “low risk” unless there appeared to be potential for bias by study design, such as the same therapist being involved in data collecting and delivering the intervention. The study by Menzies and Clarke (Menzies & Clarke, 1993) was coded as “high risk” for this reason. Additionally the “other bias” category was expanded to include consideration of researcher allegiance which is absent from the Cochrane Quality Assessment Tool (Cuijpers, 2016).

Results

The results section first outlines the range of interventions identified from the search strategy and includes a descriptive summary of these interventions. The effectiveness is then reviewed, both using a narrative summary of the findings, followed by a meta-analysis of included studies. Following this, subgroup analyses are conducted to address the aforementioned research questions and the risk of bias is formally assessed through the use of the Cochrane Quality Assessment tool

(The Cochrane Collaboration, 2011). Finally publication bias is explored and the outcome of bias- corrected analyses described.

Scope of RCTs

From 874 papers screened, 15 papers containing 13 studies were extracted for final inclusion into the review (see PRISMA flowchart, *Appendix A*). There were eight intervention studies aimed at treating anxiety disorders, four of these treatment programs were aimed at specific disorders: Separation Anxiety Disorder (Schneider et al., 2011), Darkness Phobia (Santacruz, Méndez, & Sánchez-Meca Julio, 2008), Water Phobia (Menzies & Clarke, 1993) and Obsessive Compulsive Disorder (Lewin et al., 2014); and four were more generally aimed at the treatment of anxiety problems (Cartwright-Hatton et al., 2011; Donovan & March, 2014; Hirshfeld-Becker et al., 2010; Waters, Ford, Wharton, & Cobham, 2009).

The other five studies were aimed at the prevention of the development of anxiety disorders; three were targeting Behavioural Inhibition (Anticich, Barrett, Silverman, Lacherez, & Gillies, 2013; Pahl & Barrett, 2010; Rapee, Kennedy, Ingram, Edwards, & Sweeney, 2005) whilst the other two were targeting anxiety in general (Kennedy, Rapee, & Edwards, 2009; Morgan et al., 2015, 2017).The main methods of recruitment across all studies was through local schools and advertising through various types of media, only three studies incorporated outpatient referrals from mental health clinics as part of their recruitment methodology (Hirshfeld-Becker et al., 2010; Lewin et al., 2014; Schneider et al., 2011). The amount of demographic data reported varied between studies, however on the whole of those that did report ethnicity and socioeconomic status, samples recruited were mainly from white middle-class backgrounds. Additionally, a majority of the studies (8/13) were from Australia. The reviewed studies are summarised in Table 2.

Of the 18 interventions for anxiety identified in these studies, five studies included adaptions of or variants of the same manual: “Fun Friends” (Barrett, 2007a,

2007b), “Being Brave” (Hirshfeld-Becker et al., 2010) and “Cool Little Kids”(Rapee et al., 2005), either by using these groups as active comparisons to their own intervention, or changing the format of delivery (e.g. online). The age ranges of children treated in the studies ranged from three to nine years old. In total, four of the interventions were delivered in group format (Kennedy et al., 2009; Rapee et al., 2005, 2010; Waters et al., 2009) and two of these had been adapted to be delivered online (Donovan & March, 2014; Morgan et al., 2015, 2017). The content of many of the interventions contained overlapping features; six interventions involved parent-training, five were described as Cognitive Behaviour Therapy (CBT) and nine involved the use of exposure hierarchies in their interventions. The number of sessions delivered ranged from 3-20 sessions, the mode was 10 sessions.

The scale and complexity of the interventions implemented varied widely. For instance in a treatment for water phobia, one study involved a single therapist, who conducted all three different versions of the three weekly one-to-one sessions of exposure treatments: Vicarious Exposure (VE) –watching the therapist model interactions in the water for 15 minutes followed by playing card games with therapist in room away from the pool for 15minutes;In-Vivo Exposure (IVE)- which included the same 15-minutes of card games for 15 minutes followed by 15 minutes of gradual in-vivo exposure to the water activities with therapist giving praise for activities attempted and completed; In-Vivo exposure plus Vicarious Exposure (IVVE)-this included 15 minutes of vicarious exposure followed by 15 minutes of in-vivo exposure (Menzies & Clarke, 1993).

Other interventions utilised parents as the therapists, for example combining bibliotherapy with graded in-vivo exposure-based games to address darkness phobia (Santacruz et al., 2008). Using the story of a boy addressing his fear of the dark with the help of his uncle, each of the twelve chapters are followed by a therapeutic game that progressed in difficulty; ranging from the “handkerchief game”

where the child is blindfolded and has to find a toy in his room to the more challenging "finding the noisy box" game whereby the child works through the house to find the location of the parent in the dark listening to the shaking a cereal box (Santacruz et al., 2008).

For those prevention studies that targeted Behaviour Inhibition, one study utilised a classroom-based approach in teaching children CBT strategies from the "Fun FRIENDS" program, with parents being invited to attend parent information sessions and sent weekly handouts regarding session content and homework tasks to reinforce skills learnt (Pahl & Barrett, 2010). The "Fun Friends program" encouraged "brave" behaviour (e.g. looking in people's eyes, standing up tall) and providing rewards on a chart for these behaviours to promote positive self-identity. It also incorporated teaching on affect recognition and regulation, relaxation strategies, cognitive restructuring and the creation of graded exposure hierarchies with the help of parents and teachers (with a focus on prosocial skills).

Other programmes focused more on the parent's role in maintaining anxiety; for example in the parent-education program "Cool Little Kids", the potential parent maintaining factors were targeted through providing education on the role of parent overprotection in the maintenance of anxiety and techniques to encourage and reward child's increasing independence (Morgan et al., 2015, 2017, Rapee et al., 2005, 2010).

Table 2. Characteristics of included studies.

Study		**Child demographics**				**Intervention**			**Comparator**
Authors	Country	Age range (yrs)	% female	Total *N*	Target disorder	Name	Type	Number of sessions	
Schneider, Blatter-Meunier, Herren, Adornetto, In-Albon, Lavallee (2001)	Switzerland	5-7	58.14	43	Separation anxiety disorder	'Trennungsangstprogramm für Familien' (TAFF; Separation anxiety family therapy)	Parent training + CBT (individual and family sessions).	16	Waitlist
Santacruz, Mendez & Sanchez-Meca (2008)	Spain	4- 8	47.44	78	Darkness phobia	Emotive performances (EP)	Parent training, exposure hierarchies, token economy and modelling (using in-vivo exposure).	15 (3 sessions a week for 5 weeks)	Control group
						Bibliotherapy and games (BG)	Parent training, exposure hierarchies (imaginary and in-vivo exposure).	15 (3 sessions a week for 5 weeks)	

Rapee, Kennedy, Ingram, Edwards & Sweeney (2005; 2010)	Australia	3- 5	47.94	146	Behavioural inhibition	Cool Little Kids parenting group program	Parent training, exposure hierarchies and cognitive restructuring.	6 (4 weekly 90 minute sessions, then 1 after a fortnight and final session after a month).	Waitlist
Pahl & Barrett (2010)	Australia	4- 5	51.71	263	Anxiety and Behavioural Inhibition.	Fun FRIENDS Program	CBT play-based intervention	9 (weekly 1 hour sessions). Parents invited to 3 Information sessions and given weekly handouts for home reinforcement.	Waitlist
Morgan, Rapee, Tamir, Goharpey, Salim, McLellan & Bayer (2015;2017)	Australia	4-6	50.50	433	Anxiety disorders	Cool Little Kids parenting group program adapted into online format	Parent training, exposure hierarchies and cognitive restructuring.	8 weekly online modules (30-60 minutes). Optional telephone support.	Waitlist
Menzies & Clarke (1993)	Australia	3-8	54.83	51	Water phobia	In vivo exposure plus vicarious exposure (IVVE)	Exposure-based treatment	3 (weekly for 30 minutes)	Waitlist
						Vicarious exposure (VE)	Exposure-based treatment	3 (weekly for 30 minutes)	
						In vivo exposure (IVE)	Exposure-based treatment	3 (weekly for 30 minutes)	

Lewin et al. (2014)	USA	3-8	29.03	31	OCD	Family-based Exposure/response	Behavioural parent training	12 (60 minutes delivered twice weekly over 6 weeks).	TAU
Kennedy, Rapee & Edwards (2009)	Australia	3-5	53.52	71	Anxiety disorders	Cool Little Kids parenting group program-modified	Parent training, exposure hierarchies and cognitive restructuring with an additional two sessions aimed at parent anxiety management.	8 (90min sessions) +1 (telephone follow-up call a month after completion)	Waitlist
Hirshfeld-Becker, Masek & Henin (2010)	USA	4-6	53.00	64	Anxiety disorders	Being Brave: A Program for Coping with Anxiety for Young Children and Their Parents.	Parent training + CBT	20	Waitlist
Donovan & March (2014)	Australia	3-6	53.85	52	Anxiety disorders	BRAVE-ONLINE	Online CBT with parent training	8+15-30minute telephone consultation.	Waitlist
Anticich, Barrett, Silverman, Lacherez & Gillies (2013)	Australia	4-7	55.53	488	Anxiety and Behavioural Inhibition	Fun FRIENDS	CBT	10	Waitlist
						You Can Do it	CBT	10	
Waters, Ford, Wharton & Cobham (2009)	Australia	4-8	52.5	80	Specific phobia; Social phobia,	Take ACTION (P+C)	Group CBT	10 (weekly for 60 minutes) + booster session after 8 weeks.	Waitlist

					Generalised anxiety disorder or Separation anxiety disorder	Take ACTION (PO)	Group CBT	10 (weekly for 60 minutes) + booster session after 8 weeks.	
Cartwright-Hatton, McNally, Field, Rust, Laskey & Dixon et al. (2011)	England	3-9	56.76	74	Anxiety disorder	Timid to Tiger	Parent training + CBT	10 (2 hour sessions)	Waitlist

a Treatment aimed at children but included Learning Adventure Workbook for consolidation skills to be used at home with parents. Parents invited to two parent sessions.

What works?

It is important to highlight that only six studies conducted follow-up measures beyond the immediate post-treatment period for both the control and intervention groups, the most typical reason cited for this was due to ethics regarding withholding treatment from the control group participants (see Table 1, *Appendix C*). Of those that implemented a follow-up assessment, some found the main effects of the treatment on anxiety measures were no longer significant compared to the control group (Cartwright-Hatton et al., 2011; Rapee et al., 2005, 2010). However, two studies found no differences between treatment and control groups at post-intervention but did find significant improvements in the intervention group at follow up (Anticich et al., 2013; Pahl & Barrett, 2010). The rest of the studies showed the treatment condition maintained superior outcome to control at follow-up.

Main meta-analysis findings

The results of the meta-analysis of the intervention studies suggests a large overall effect (*d*= .80; 95% *CI*=0.51-1.08). The funnel plot of the meta-analysis shows there are no outlier studies: the 95% confidence intervals of the individual studies overlapped with the pooled effect size confidence interval (see Cuijpers, 2016). Whilst the effect sizes of the interventions clearly differ, the studies with larger sample sizes (Anticich et al., 2013; Morgan et al., 2015, 2017; Pahl & Barrett, 2010; Rapee et al., 2005, 2010) cluster to the left-hand side of the effect size range and closer to the point of 0. The heterogeneity of the effects reported by these studies is considered to be high, I^2 =80% (Higgins, Thompson, Deeks, & Altman, 2003, as cited in Cuijpers, 2016) and thus further subgroup analyses is warranted.

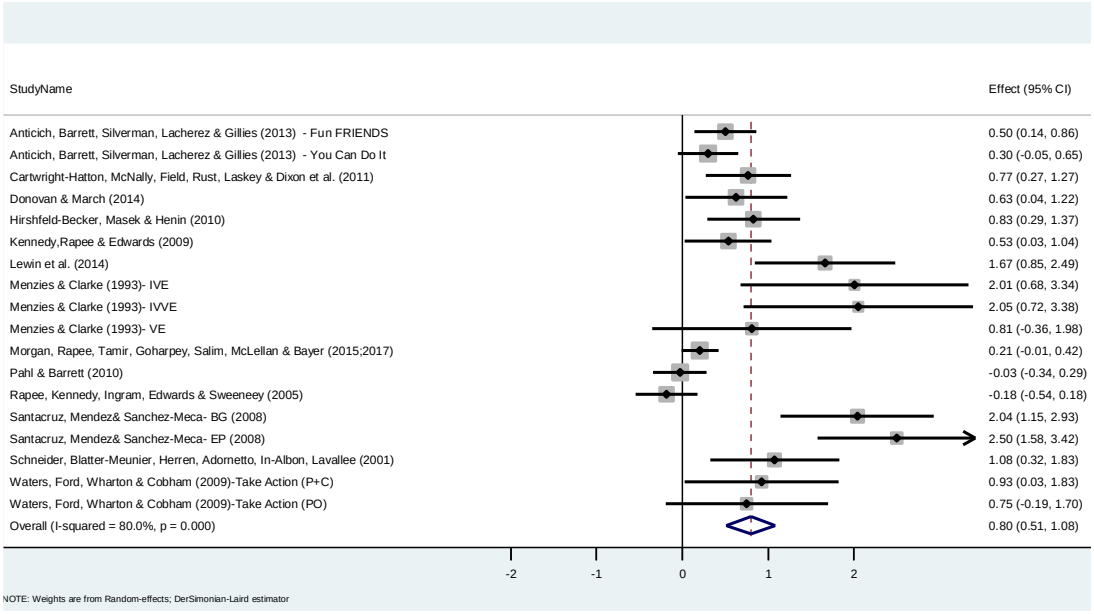


Figure 3. Figure showing forest plot of meta-analysis.

Subgroup analysis I: Prevention vs Intervention studies

When prevention and intervention study designs were compared, the heterogeneity of the studies was reduced to more moderate levels, I^2= 54.3% (*p*=.05) and 55.4% (*p*=.01) respectively. There was a significant difference in the effect sizes between these two groups of studies (*p*=<.001), with the effect size for prevention studies showing a small effect size (*d*=.20, 95% CI= 0.00- 0.40) compared to the intervention studies, which showed a large effect size (*d*=1.24, 95% *CI*= 0.89-1.59). It is also important to note that the prevention studies subgroup contains all of the RCTs with large sample sizes (*N* > 71-488).

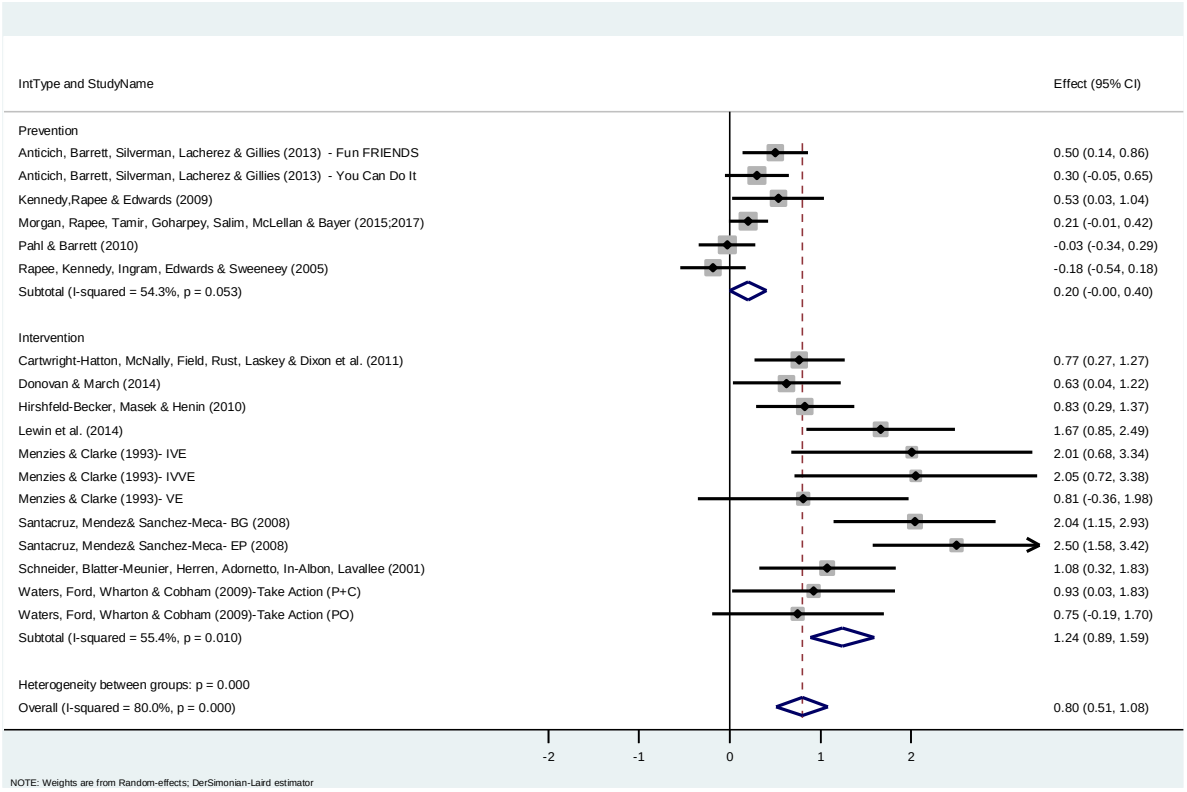


Figure 4. Forest plot of metanalysis divided by intervention and intervention subgroups.

Subgroup analysis II: Involvement of parents in the treatment

The heterogentity of studies in each group ranged from moderate for the child only and parent only subgroup groups, I^2=65.6% (*p*=.02) and I^2=63.5% (*p*=.02), to high for the parent and child subgroup, I^2=88.1%, (*p*<.001). The effect sizes for the parent and child subgroup (*d*=1.24, 95% *CI*=0.51-1.96) and child only subgroup (*d*=.81, 95% *CI*= 0.29-1.34) indicate large average effects, with the parent only subgroup having a small effect size (*d*= .37, 95% *CI*= 0.07-0.67), a significant difference between subgroups (*p*<.001). Given the small number of studies in each subgroup and high levels of study heterogenity, caution should be taken in interpreting these results.

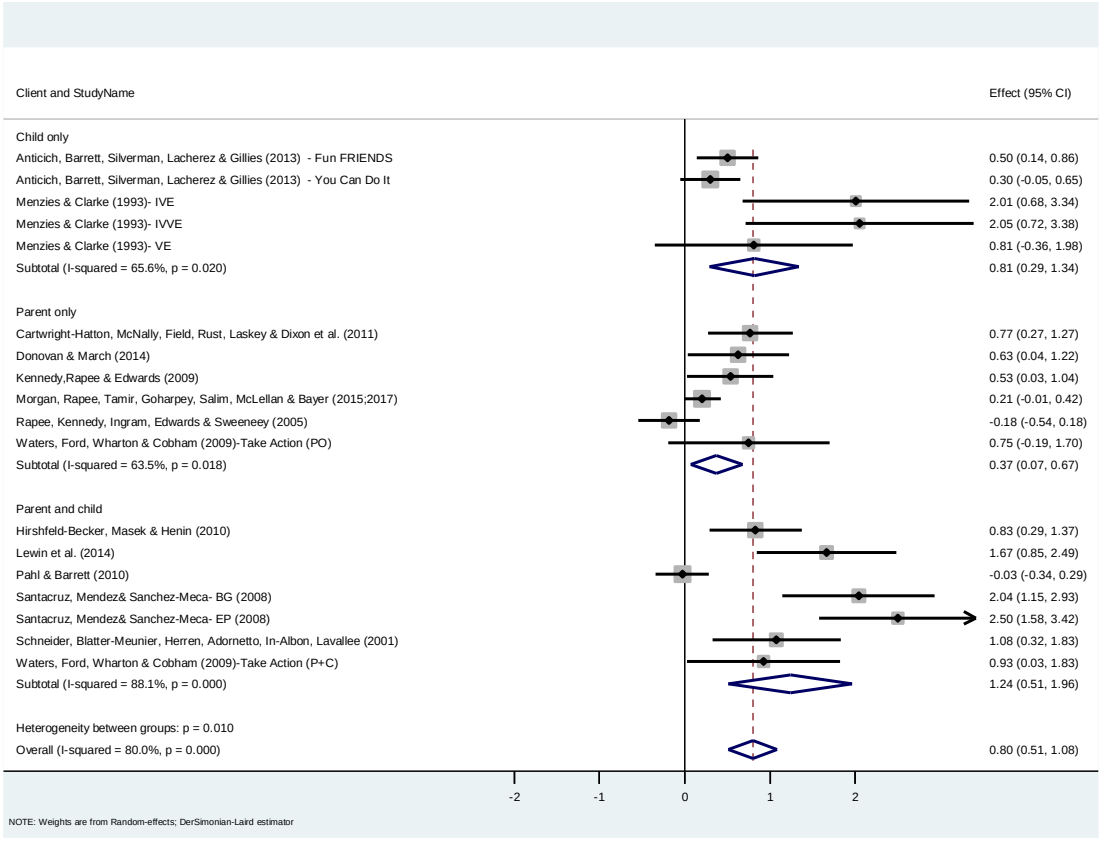


Figure 5. Forest plot of metanalysis divided by client group.

Subgroup analysis III: CBT vs other approaches

Studies were categorised as CBT if they described the use of *both* cognitive restructuring and behavioural techniques (e.g. exposure) in the treatment (or the study explicitly indicated that the treatment was CBT). Only three studies utilised interventions that did not meet that criteria and instead used treatments that adopted primarily a behavioural approach, mainly focusing on elements of exposure only (Lewin et al., 2014; Menzies & Clarke, 1993; Santacruz et al., 2008).

The CBT subgroup shows a moderate effect size for CBT interventions (*d*=.42, *p*<.001, 95% *CI* :0.21-0.63), and reduced the heteogentiy of studies to a moderate level for the CBT subgroup (I^2=61.9%, *p*<.001). The Other subgroup produced a large effect size and did not meet significance on tests of heterogenity (*d*=1.88, p<.001, 95% *CI*:1.44-2.31, I^2= 8.1%, *p*=.36). However, the limited heterogeneity in this instance may reflect the small number of studies in this

subgroup, leading to insufficient power. Furthermore, this subgroup included only three separate studies, and so the averaged effect may not be reliable.

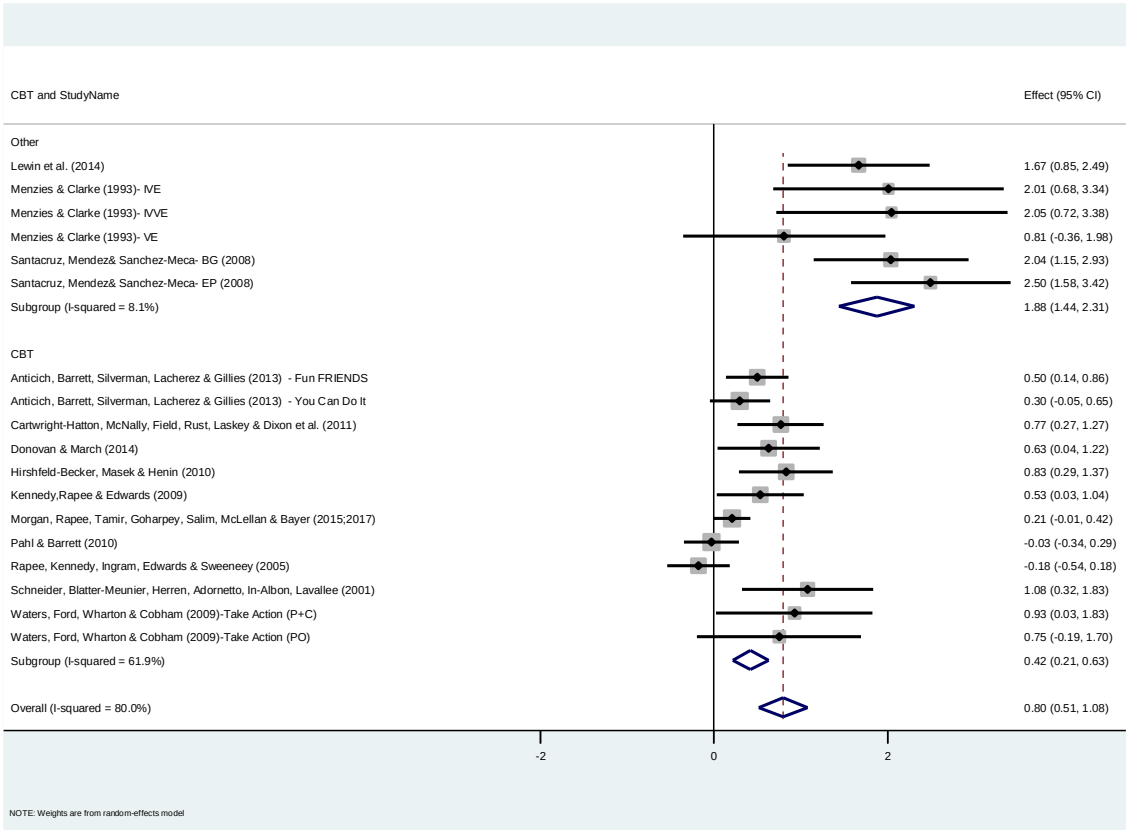


Figure 6. Forest plot of metanalysis divided by interverntions using CBT techniques and those without.

Risk of bias in included studies

Of the 13 studies extracted and assessed using the Cochrane Quality Assessment Tool (The Cochrane Collaboration, 2011), six studies scored “high risk” in at least one assessment category (Anticich et al., 2013; Menzies & Clarke, 1993; Pahl & Barrett, 2010; Rapee et al., 2005, 2010; Santacruz et al., 2008; Schneider et al., 2011; Waters et al., 2009) and only one study (Lewin et al., 2014) scored “low risk” for all risk items. Lewin and colleagues (2014) examined the efficacy of family-based Exposure/Response Prevention (E/RP) therapy in treating early onset OCD for 31 children aged three to eight years. This study adapted a manualised treatment of childhood OCD (Freeman et al., 2014) which targets OCD symptoms

and accommodation of OCD by family members. This intensive treatment of twice-weekly hour sessions for six weeks required at least one consistent parent present to attend with the child for every session, using psychoeducation, development of rewards program, differential reinforcement and E/RP. They found that compared to the treatment as usual condition, E/RP treatment produced superior outcomes on reduction of OCD symptoms (*d*=1.69) with treatment gains maintained at a three month follow-up.

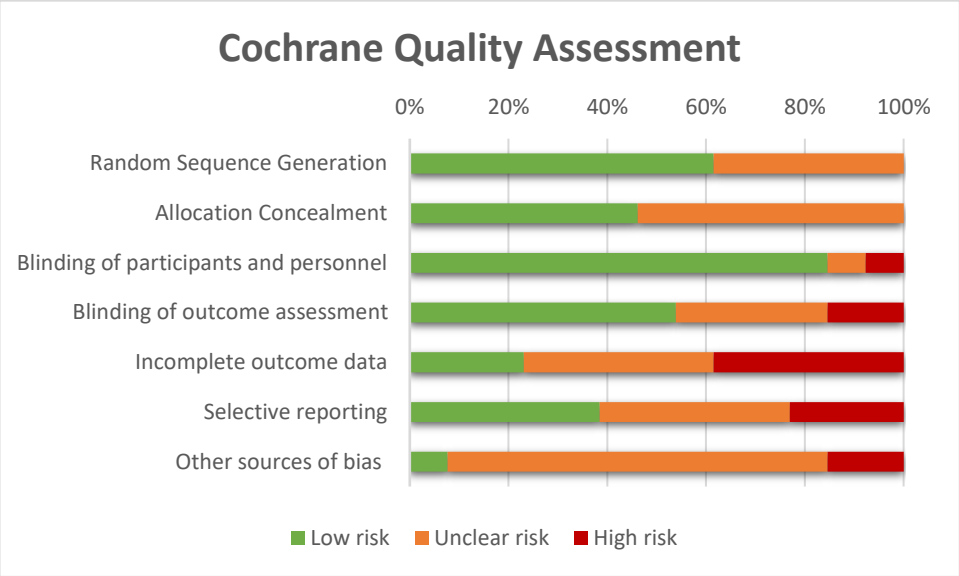


Figure 7. Summary of risk of bias: each risk item is shown as percentages across all studies.

Selection bias

Most studies reported adequate "low risk" randomisation sequencing methods (Cartwright-Hatton et al., 2011; Donovan & March, 2014; Hirshfeld-Becker et al., 2010; Kennedy et al., 2009; Lewin et al., 2014; Morgan et al., 2015, 2017, Rapee et al., 2005, 2010; Schneider et al., 2011), using a-priori randomisation lists, coin toss and computer-based randomisation. The rest of the studies failed to report

randomisation methods and therefore were labelled as “unclear risk” as it was not possible to ascertain potential bias. Fewer studies reported on allocation concealment, with those that did report on this scoring “low risk”(Cartwright-Hatton et al., 2011; Donovan & March, 2014; Hirshfeld-Becker et al., 2010; Kennedy et al., 2009; Lewin et al., 2014; Morgan et al., 2015, 2017). The most common method to preserve allocation concealment was through using an independent person for randomisation, although some randomisation methods prevented selection bias, e.g. coin toss.

Performance bias

As mentioned previously, as all the reviewed trials examined psychotherapeutic interventions, it was generally not possible for participants to be blind to the treatment condition they were in nor was this a meaningful criterion for interventions with a waitlist control (Cartwright-Hatton et al., 2011; Donovan & March, 2014; Hirshfeld-Becker et al., 2010; Kennedy et al., 2009; Menzies & Clarke, 1993; Morgan et al., 2015, 2017; Pahl & Barrett, 2010; Rapee et al., 2005, 2010; Schneider et al., 2011; Waters et al., 2009). Therefore this risk item was modified to consider the potential performance bias from the therapist delivering the intervention, in which one study scored “high risk” due to using the same therapist to deliver all three conditions and collect outcome measures based on own observations (Menzies & Clarke, 1993) and one scored “unclear risk” as the primary author delivered the treatment in both conditions, although different researchers collected the outcome measures (Waters et al., 2009).

Detection bias

The blinding of outcome assessments was coded to assess whether knowledge from the assessors about allocated interventions could have produced detection bias in the findings of the study. Only two studies were deemed “high risk” (Menzies & Clarke, 1993; Waters et al., 2009), whereby the therapist also collected

the outcome measurements. For four studies there was insufficient information provided to assess potential risk of bias (Anticich et al., 2013; Kennedy et al., 2009; Pahl & Barrett, 2010; Santacruz et al., 2008).

Attrition bias

This criteria examined the completeness of outcome data reported, including attrition rates and the method of handling incomplete outcome data in the study that could have biased the outcome and conclusions from the study. Only three studies scored “low risk” of bias on this item (Hirshfeld-Becker et al., 2010; Lewin et al., 2014; Schneider et al., 2011). Studies scored “high risk” for: not reporting reasons for attrition or not using Intention to Treat analyses (Menzies & Clarke, 1993; Santacruz et al., 2008); not conducting significance tests for the attrition when there were different attrition rates for each arm with no reasons for attrition reported by each group (Rapee et al., 2005, 2010); having large attrition rates and unclear reporting from which arm (Anticich et al., 2013; Pahl & Barrett, 2010). Studies were rated “unclear risk” when there was the potential presence of reporting bias but other mitigating factors meant that overall level of bias was hard to ascertain; for example low or balanced numbers of attrition, or had administered significance testing for completers vs non-completers and adjusted for this in analysis.

Reporting bias

Three studies were judged to be “high risk” of selective outcome reporting due to incomplete reporting (Anticich et al., 2013; Santacruz et al., 2008; Schneider et al., 2011). The majority of studies were rated low risk; of those, five studies reported clear primary and secondary outcomes (Donovan & March, 2014; Hirshfeld-Becker et al., 2010; Lewin et al., 2014; Morgan et al., 2015, 2017; Pahl & Barrett, 2010) and only one had a published protocol (Morgan et al., 2015, 2017). The remaining studies received an “unclear risk” rating which included studies that

did not publish a protocol or did not have pre-specified primary and secondary outcomes, but did provide complete reporting on the outcome measures.

Other bias

This criterion aimed to assess potential bias caused by researcher allegiance to the intervention in question and other areas of bias. Only one study was rated as "low risk" of bias because there were no indications that the authors would profit from the success of the intervention (i.e. developed the manual being investigated) or was being funded for such purpose (Lewin et al., 2014). Two studies were rated as "high risk bias", because both studies had some clear interest in the success of the treatment being implemented plus other concerns including: non-adherence to treatment protocol (Anticich et al., 2013); the addition of a self-selecting sample in the recruitment method and inconsistences in data reported regarding demographics and outcome data between two papers of the same study (Rapee et al., 2005, 2010). All remaining studies were rated "unclear bias" due to concerns regarding researcher allegiance as a result of funding or vested interest in treatment success for personal gain (i.e. success of treatment manual created).

Publication bias

The funnel plot of studies illustrating reduction of anxiety symptoms displayed an asymmetrical figure which denoted potential publication bias, with a lack of smaller sample studies favouring control as would be expected by chance (Cuijpers, 2016). This was consistent with the results of the Egger's test (Egger, Smith, Schneider, & Minder, 1997) that detected possible publication bias (t=5.45, p<.001). As a result of possible publication bias, the trim and fill method (Duval & Tweedie, 2000) was used. This adjusted model (Figure 8) significantly reduced the estimated effect to a small effect size (d=.40, 95% *CI*: 0.11-0.66, p=.01).

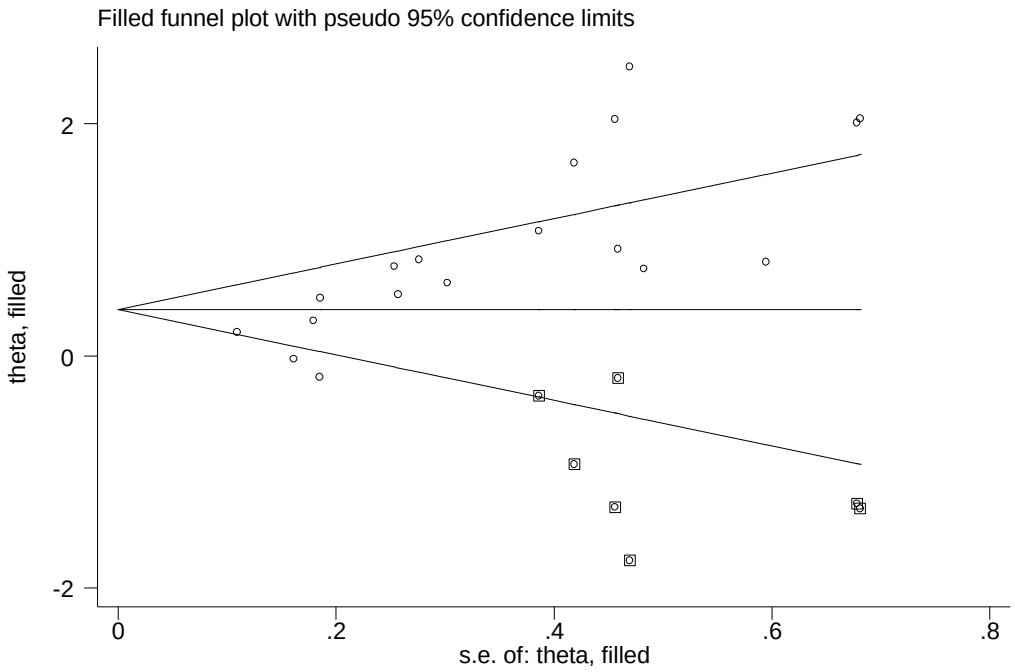


Figure 8. Funnel plot using trim and fill method. Square boxes indicate “filled” studies.

Whilst this could present publication bias it is also important to consider the small number of studies included in this review and heterogeneity of study design (most notably between the prevention and intervention studies). It is a possibility that smaller studies have targeted high risk populations in their sample selection and may thus be targeting different populations. Therefore tests for publication bias were run separately for intervention and prevention studies.

For the intervention studies subgroup, the Egger’s test (Egger et al., 1997) indicated the presence of publication bias (t=2.51, p=.03). Nevertheless, the adjusted analysis using the trim and fill method (Duval & Tweedie, 2000) still produced a large effect size for the intervention studies (d=1.09, 95% *CI*:0.71-1.46, p<.001). In contrast the prevention studies subgroup did not meet significance for Egger’s test (Egger et al., 1997) suggesting absence of publication bias.

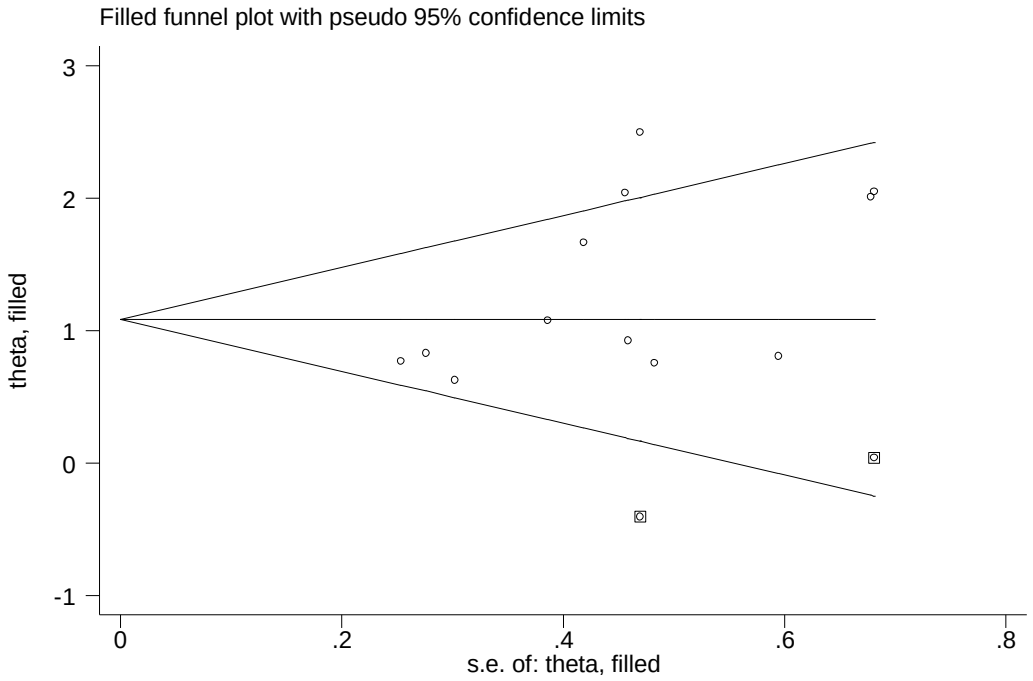


Figure 9. Funnel plot of studies from the Intervention subgroup using trim and fill method. Squares indicate “filled” studies.

Discussion

Summary of main results

The search strategy identified a small number of included studies (*N*=13) which provided data for 18 interventions for preschool anxiety. The results of the review included a wide range of studies, which appeared to be best conceptualised as two subgroups: intervention and prevention studies. The data suggested that the intervention studies evaluating treatments for preschool anxiety were promising, even the more conservative adjusted meta-analysis produced a large effect size (*d*=1.09), however large confidence intervals suggested heterogeneity in these findings. The results from the prevention studies subgroup appeared less promising, despite containing the largest trials, the meta-analysis found a small effect size that just fell short of significance (*p*=.05). The quality assessment tool also highlighted some methodological shortfalls in the current studies.

What are the range of interventions for treating childhood anxiety?

This review highlighted a wide range of types of interventions for this age group that varied in target disorder, content and involvement of parents. Most interventions included the use of exposure techniques (in some format), with twelve studies using this in combination with cognitive restructuring. The use of the parent in these interventions varied, some interventions targeted parental anxiety directly in their interventions through cognitive restructuring (e.g. Rapee et al., 2005, 2010) and some trained the parent as a therapist to the child (e.g. Santacruz et al., 2008). It is important to note that all interventions , with one exception (Menzies & Clarke, 1993), involved parents in some capacity, from the more minimal optional attendance at a parents' information evening (e.g. Pahl & Barrett, 2010) to offering a combination of individual and joint sessions (e.g. Schneider et al., 2011). The modality of delivery was also heterogeneous, with interventions being offered in the form of individual sessions (Menzies & Clarke, 1993), family sessions (e.g. Schneider et al., 2011), groups (e.g. Waters et al., 2009) and online (e.g. Donovan & March, 2014).

How effective are interventions at treating and preventing anxiety disorders?

On the whole the evidence supported psychotherapeutic interventions targeting preschool anxiety, consistent with a previous narrative review of the literature (Luby, 2013) which concluded that developmentally adapted forms of behavioural and CBT forms of treatment for this age group are both feasible and beneficial. Whilst the overarching meta-analysis showed a large effect in favour of treatment (*d*=.8), large confidence intervals and heterogeneity in the model warranted further analysis. The size of the heterogeneity from the different studies raised questions as to whether the samples drawn from these studies were representative of a common population, thus limiting the extent of conclusions drawn from the data as presented in its pooled form. This was partially supported by the subgroup analysis (see Figure 4) which showed the division of studies into

prevention and intervention subgroups reduced the amount of heterogeneity in the model.

The subgroup analysis showed that small and unreliable treatment effects were observed in the prevention studies subgroup (which became no longer statistically significant) while the effect size for the intervention subgroup was large. This division makes sense: cases recruited in the intervention group would have been diagnosed with an anxiety disorder, thus represent a more homogeneous population with greater room for consistent improvement following intervention. However the true difference between the effect sizes of these subgroups may not be accurately presented from this data due to the imbalance of large trials in the analysis, with the prevention subgroup containing all of the studies with the larger sample. Thus it can be inferred that whilst the prevention subgroup may present relatively accurate effect size estimate, there may be less confidence in the accuracy of the magnitude of the intervention subgroup effect size due to the smaller sample sizes in these studies. Nevertheless, the large overall effect for intervention/treatment studies suggests that bias is unlikely to reduce the effect to clinically trivial levels. Furthermore, these findings are in agreement with other meta-analytic reviews supporting the effectiveness of psychotherapeutic interventions targeting anxiety in later childhood (Ewing et al., 2015; Hirshfeld-Becker, Micco, Simoes, et al., 2008; James et al., 2015).

Are these interventions more effective if utilising parents in their intervention?

Previous reviews of the literature have recommended that parents should be centrally involved in all interventions offered for preschool anxiety, whilst also ensuring the child is the main target of the intervention and not solely targeting parental anxiety (Luby, 2013). Only interventions that mainly targeted preschool child anxiety were used in this review, thus comparisons of interventions targeting parental anxiety only are outside of the remit of this review. Comparisons were

conducted of interventions that employed different levels of parent involvement, which found a significant difference between the three different levels of parental involvement in studies (parents only, parent and child, predominantly child only). The interventions that used parents only to deliver the intervention received a small effect size *(d*=.37), compared to the large effect size seen for the child only subgroup (*d*=.81) and the parent and child subgroup (*d*=1.24). These results seem to favour direct input with child alongside parental involvement, however it is important to note the large confidence intervals and heterogeneity indicated in this model warrants caution. It is possible that other factors not controlled for in the current analysis could have affected the results, such as the number of sessions offered, content of sessions, modality of delivery (e.g. groups, etc).

Are interventions using CBT techniques more effective?

As a result of the different levels of clarity and detail in the descriptions of the treatment procedures the included studies, a more crude analysis was conducted in the present review, whereby "CBT" interventions (that were deemed to have included an element of cognitive restructuring and behavioural exposure) were compared against "other" interventions. Remarkably there were only six interventions that did not meet the "CBT" criteria but utilised behavioural approaches based on different forms of exposure. For the CBT subgroup there was a moderate effect size (*d*=.42) for CBT interventions with relatively small confidence intervals.

Other meta-analytic reviews for children have found odds ratio of 7.85 (95% *CI*: 5.31 -11.60)(James et al., 2015) and 9.15 (95% *CI*: 6.05 –13.87)(Ewing et al., 2015). The odds ratio for the CBT subgroup was 2.14, comparatively smaller than the aforementioned reviews. A possible reason for this is the inclusion of prevention studies within this subgroup analysis, whereby there would be an expected smaller reduction in symptoms. Additionally, unlike the other reviews, the present review

focuses solely on the preschool age, so therefore it could be possible that even for age-adapted CBT interventions the developmental differences may impact on the efficacy of this approach. Due to the small sample sizes and number of studies that were in the “other” subgroup a reliable estimation of an effect size was not possible.

Limitations

The heterogeneity in the meta-analysis was high, even when conducting subgroup analyses. Studies varied in terms of recruitment methodology, measures used and disorders targeted. Subgroup analyses were conducted to explore whether intervention type, level of parental involvement and content of intervention which accounted for only small levels of heterogeneity. Previous reviews have suggested meta-analytic reviews should focus on efforts to guide anxiety prevention choice (Bennett et al., 2015) however with the limited number of RCTs in this area attempts to perform specific analyses on more specific variables of interest (e.g. disorder specific vs transdiagnostic CBT) would not be possible.

Similarly as only thirteen studies were included, the meta-analysis may be vulnerable to bias, either from methodological shortfalls identified from the quality assessment or from shared commonalities in studies and missing studies. Notably, most of the trials, including the majority of the larger trials piloted, were from Australia, and five of those studies contained primary research investigators that had investigated the efficacy of three intervention manuals they had created, thus potentially meaning some of the larger trails could be biased by researcher allegiance. The quality assessment tool highlighted that less than half of the studies scored “low risk” for incomplete outcome data and selective reporting which, combined with the high risk of publication bias for the intervention studies, suggests that caution should be taken in interpreting the results of the meta-analysis.

It is also important to recognise that pragmatic decisions made when the analysis was conducted may have also affected the effect sizes. Due to the small

number of intervention studies investigated, and to reduce heterogeneity, active control groups that offered a specific intervention were considered as separate interventions and compared to the control group (e.g. Anticich et al., 2013). As a result of this only one study used treatment as usual as a control group and the rest of the interventions utilised a waitlist control group. It is therefore possible that this also may result in an inflation of the observed effect sizes.

Additionally, the meta-analysis used post-intervention data only due to an absence of follow-up data for control groups in some study designs (often due to ethical constraints that did not permit deprivation of treatment for such time). For the sake of uniformity, the earliest time points were extracted for all papers. For some studies this meant that effect sizes were reduced compared to what was reported in the original paper (e.g. Cartwright-Hatton et al., 2011). It also may mean for the prevention subgroup of studies that the potential impact of the treatments on the prevention of preschool anxiety disorders may not be ideally captured, as the impact of the intervention may be predicted to have longer-term positive implications (e.g. Pahl & Barrett, 2010).

Future research

Whilst it is remarkable that researchers in Australia are attempting to bridge the existing gap in the literature for preschool anxiety interventions (Chavira et al., 2004), it is important that other countries and research teams also prioritise and replicate this work to ensure ecological validity in other countries and health care systems.

The scope of this review was limited to randomised controlled trials as the review sought to address the effectiveness of preschool interventions on anxiety. As a result this may mean that broadening the search strategy to include controlled trials may produce a larger range of interventions and identify approaches outside of behavioural and CBT approaches in treating child anxiety in this field.

Finally, in this meta-analysis prevention studies included the highest sample sizes and used non-clinical samples that highlighted a small effect size that only just missed significance (p=.05). Future reviews seeking to examine the longer-term impact of these interventions would provide more illuminating insight into the value of these interventions. Further work aimed at improving prevention outcomes would be valuable in the future. Additionally, future research using intervention studies that target clinical samples would benefit from using larger sample sizes to improve accuracy of the effect size for preschool anxiety interventions.

Conclusions

This is the first attempt to synthesise the literature of randomised-controlled trials seeking to treat anxiety in the preschool age group. Whilst the heterogeneity of the included studies leads to some concern regarding the accuracy of the effect sizes reported, it nonetheless points to clear trends in the current literature that support the feasibility of using psychological interventions as an effective first line of treatment in targeting preschool anxiety. The results have illustrated that interventions provided in various formats can be effective in reducing preschool anxiety, and provided support for developmentally-adapted CBT-based interventions which mirrors the findings of meta-analytic reviews of older children.

Whilst these findings show promise for the provision of treatment in a clinical setting, with larger effects seen in the reduction of anxiety for preschool children already diagnosed with an anxiety disorder, there were also promising effects of providing preventive interventions to preschool children showing signs of anxiety or high levels of behavioural inhibition. Thus providing further support for the importance of early intervention as an important preventative measure, with wider implications for public health policy and the education sector. This review also

highlighted the lack of attention to preschool anxiety on research agendas outside of Australia, where all of the largest randomised controlled trials were conducted.

References

Anticich, S. A. J., Barrett, P. M., Silverman, W., Lacherez, P., & Gillies, R. (2013). The prevention of childhood anxiety and promotion of resilience among preschool-aged children: A universal school based trial. *Advances in School Mental Health Promotion*, *6*(2), 93–121. http://doi.org/10.1080/1754730X.2013.784616

Barrett, P. M. (2007a). *Fun Friends Family Learning Adventure: Resilience Building Activities for 4, 5, & 6 year old Children*. Brisbane, Australia: Fun Friends Publishing.

Barrett, P. M. (2007b). *Fun Friends The Teaching and Training Manual for Group Leaders.* Brisbane, Australia: Fun Friends Publishing.

Bennett, K., Manassis, K., Duda, S., Bagnell, A., Bernstein, G. A., Garland, E. J., … Wilansky, P. (2015). Preventing Child and Adolescent Anxiety Disorders: Overview of Systematic Reviews. *Depression and Anxiety*, *32*(12), 909–918. http://doi.org/10.1002/da.22400

Bennett, K., Manassis, K., Walter, S., Cheung, A., Wilansky-Traynor, P., Diaz-Granados, N., … Wood, J. (2013). Cognitive behavioral therapy age effects in child and adolescent anxiety: An individual patient data metaanalysis. *Depression and Anxiety*, *30*(9), 829–841. http://doi.org/10.1002/da.22099.

Bittner, A., Egger, H. L., Erkanli, A., Jane Costello, E., Foley, D. L., & Angold, A. (2007). What do childhood anxiety disorders predict? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *48*(12), 1174–1183. http://doi.org/10.1111/j.1469-7610.2007.01812.x

Bufferd, S. J., Dougherty, L. R., Carlson, G. A., & Klein, D. N. (2011). Parent-reported mental health in preschoolers: findings using a diagnostic interview.

Comprehensive Psychiatry, *52*(4), 359–369. http://doi.org/https://doi.org/10.1016/j.comppsych.2010.08.006

Cartwright-Hatton, S., McNally, D., Field, A. P., Rust, S., Laskey, B., Dixon, C., … Woodham, A. (2011). A new parenting-based group intervention for young anxious children: Results of a randomized controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *50*(3), 242–251. http://doi.org/10.1016/j.jaac.2010.12.015

Chavira, D. A., Stein, M. B., Bailey, K., & Stein, M. T. (2004). Child anxiety in primary care: Prevalent but untreated. *Depression and Anxiety*, *20*(4), 155–164. http://doi.org/10.1002/da.20039

Cuijpers, P. (2016). *Meta-analyses in mental health research . A practical guide*.

Danzig, A. P., Bufferd, S. J., Dougherty, L. R., Carlson, G. A., Olino, T. M., & Klein, D. N. (2013). Longitudinal associations between preschool psychopathology and school-age peer functioning. *Child Psychiatry and Human Development*, *44*(5), 621–632. http://doi.org/10.1007/s10578-012-0356-4

Dersimonian, R., & Laird, N. (1986). Meta-Analysis in Clinical Trials*. *Statistics in Medicine*, *188*, 177–188. http://doi.org/10.1016/0197-2456(86)90046-2

Donovan, C. L., & March, S. (2014). Online CBT for preschool anxiety disorders: A randomised control trial. *Behaviour Research and Therapy*, *58*, 24–35. http://doi.org/10.1016/j.brat.2014.05.001

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. http://doi.org/10.1111/j.0006-341X.2000.00455.x

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child*

Psychology and Psychiatry and Allied Disciplines, *47*(3–4), 313–337. http://doi.org/10.1111/j.1469-7610.2006.01618.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. http://doi.org/10.1136/bmj.315.7109.629

Ewing, D. L., Monsen, J. J., Thompson, E. J., & Cartwright-hatton, S. (2015). Europe PMC Funders Group Europe PMC Funders Author Manuscripts A Meta-Analysis of Transdiagnostic Cognitive Behavioural Therapy in the Treatment of Child and Young Person Anxiety Disorders, *43*(5), 562–577. http://doi.org/10.1017/S1352465813001094.A

Franz, L., Angold, A., Copeland, W., Costello, E. J., Towe-Goodman, N., & Egger, H. (2013). Preschool Anxiety Disorders in Pediatric Primary Care: Prevalence and Comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(12), 1294–1303.e1. http://doi.org/https://doi.org/10.1016/j.jaac.2013.09.008

Freeman, J., Sapyta, J., Garcia, A., Compton, S., Khanna, M., Flessner, C., … Franklin, M. (2014). Family-Based Treatment of Early Childhood Obsessive-Compulsive Disorder. *JAMA Psychiatry*, *71*(6), 689. http://doi.org/10.1001/jamapsychiatry.2014.170

Gadow, K. D., Sprafkin, J., & Nolan, E. E. (2001). DSM-IV Symptoms in Community and Clinic Preschool Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(12), 1383–1392. http://doi.org/10.1097/00004583-200112000-00008

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ : British Medical Journal*, *327*(7414), 557–560. http://doi.org/10.1136/bmj.327.7414.557

Hirshfeld-Becker, D. R., Masek, B., Henin, A., Blakely, L. R., Pollock-Wurman, R. A., McQuade, J., … Biederman, J. (2010). Cognitive behavioral therapy for 4- to 7-year-old children with anxiety disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *78*(4), 498–510. http://doi.org/10.1037/a0019055

Hirshfeld-Becker, D. R., Micco, J. A., Simoes, N. A., & Henin, A. (2008). High risk studies and developmental antecedents of anxiety disorders. *American Journal of Medical Genetics, Part C: Seminars in Medical Genetics*, *148*(2), 99–117. http://doi.org/10.1002/ajmg.c.30170

Hirshfeld-Becker, D. R., Micco, J., Henin, A., Bloomfield, A., Biederman, J., & Rosenbaum, J. (2008). Behavioral inhibition. *Depression and Anxiety*, *25*(4), 357–367. http://doi.org/10.1002/da.20490

James, A. C., James, G., Cowdrey, F. A., Soler, A., Choke, A., Ac, J., … Choke, A. (2015). Cognitive behavioural therapy for anxiety disorders in children and adolescents (Review). *The Cochrane Database of Systematic Reviews*, *2*(2), 1–108. http://doi.org/10.1002/14651858.CD004690

Katrak, P., Bialocerkowski, A. E., Massy-Westropp, N., Kumar, V. S., & Grimmer, K. A. (2004). A systematic review of the content of critical appraisal tools. *BMC Medical Research Methodology*, *4*(1), 22. http://doi.org/10.1186/1471-2288-4-22

Kennedy, S. J., Rapee, R. M., & Edwards, S. L. (2009). A Selective Intervention Program for Inhibited Preschool-Aged Children of Parents With an Anxiety Disorder: Effects on Current Anxiety Disorders and Temperament. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(6), 602–609. http://doi.org/10.1097/CHI.0b013e31819f6fa9

Lewin, A. B., Park, J. M., Jones, A. M., Crawford, E. A., DeNadai, A. S., Menzel, J.,

… Storch, E. A. (2014). Family-based exposure and response prevention therapy for preschool-aged children with obsessive-compulsive disorder: A pilot randomized controlled trial. *Behaviour Research and Therapy*, *56*(1), 30–38. http://doi.org/10.1016/j.brat.2014.02.001

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis. Applied social research methods series* (Volume 49). California: Thousand Oaks, Sage.

Luby, J. L. (2013). Treatment of anxiety and depression in the preschool period. *Journal of the American Academy of Child and Adolescent Psychiatry*, *52*(4), 346–358. http://doi.org/10.1016/j.jaac.2013.01.011

McLeod, B. D., Wood, J. J., & Weisz, J. R. (2007). Examining the association between parenting and childhood anxiety: A meta-analysis. *Clinical Psychology Review*, *27*(2), 155–172. http://doi.org/10.1016/j.cpr.2006.09.002

Menzies, R. G., & Clarke, J. C. (1993). A comparison of in vivo and vicarious exposure in the treatment of childhood water phobia. *Behaviour Research and Therapy*, *31*(1), 9–15. http://doi.org/10.1016/0005-7967(93)90037-U

Morgan, A. J., Rapee, R. M., Salim, A., Goharpey, N., Tamir, E., McLellan, L. F., & Bayer, J. K. (2017). Internet-Delivered Parenting Program for Prevention and Early Intervention of Anxiety Problems in Young Children: Randomized Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(5), 417–425. http://doi.org/10.1016/j.jaac.2017.02.010

Morgan, A. J., Rapee, R. M., Tamir, E., Goharpey, N., Salim, A., McLellan, L. F., & Bayer, J. K. (2015). Preventing anxiety problems in children with Cool Little Kids Online: study protocol for a randomised controlled trial. *Trials*, *16*(1), 507. http://doi.org/10.1186/s13063-015-1022-5

Muris, P., van Brakel, A. M. L., Arntz, A., & Schouten, E. (2011). Behavioral

Inhibition as a Risk Factor for the Development of Childhood Anxiety Disorders: A Longitudinal Study. *Journal of Child and Family Studies*, *20*(2), 157–170. http://doi.org/10.1007/s10826-010-9365-8

National Institute for Health and Care Excellence. (2012). Social and emotional wellbeing: early years Public health guideline. Retrieved December 9, 2017, from https://www.nice.org.uk/guidance/ph40

Pahl, K. M., & Barrett, P. M. (2010). Preventing anxiety and promoting social and emotional strength in preschool children: A universal evaluation of the fun FRIENDS program. *Advances in School Mental Health Promotion*, *3*(3), 14–25. http://doi.org/10.1080/1754730X.2010.9715683

Rapee, R. M., Kennedy, S., Ingram, M., Edwards, S., & Sweeney, L. (2005). Prevention and Early Intervention of Anxiety Disorders in Inhibited Preschool Children. *Journal of Consulting and Clinical Psychology*, *73*(3), 488–497. http://doi.org/10.1037/0022-006X.73.3.488

Rapee, R. M., Kennedy, S. J., Ingram, M., Edwards, S. L., & Sweeney, L. (2010). Altering the trajectory of anxiety in at-risk young children. *American Journal of Psychiatry*, *167*(12), 1518–1525. http://doi.org/10.1176/appi.ajp.2010.09111619

Santacruz, I., Méndez, F. J., & Sánchez-Meca Julio. (2008). Play Therapy Applied by Parents for Children with Darkness Phobia: Comparison of Two Programmes. *Child & Family Behavior Therapy*, *28*(1), 19–35. http://doi.org/10.1300/J019v28n01

Schneider, S., Blatter-Meunier, J., Herren, C., Adornetto, C., In-Albon, T., & Lavallee, K. (2011). Disorder-specific cognitive-behavioral therapy for separation anxiety disorder in young children: A randomized waiting-list-controlled trial. *Psychotherapy and Psychosomatics*, *80*(4), 206–215.

http://doi.org/10.1159/000323444

Sterba, S. K., Prinstein, M. J., & Cox, M. J. (2007). Trajectories of internalizing problems across childhood: Heterogeneity, external validity, and gender differences. *Development and Psychopathology*, *19*(2), 345–366. http://doi.org/10.1017/S0954579407070174

The Cochrane Collaboration. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].* (H. JPT & S. Green, Eds.).

Towe-Goodman, N. R., Franz, L., Copeland, W., Angold, A., & Egger, H. (2014). Perceived family impact of preschool anxiety disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(4), 437–446. http://doi.org/10.1016/j.jaac.2013.12.017

Waters, A. M., Ford, L. A., Wharton, T. A., & Cobham, V. E. (2009). Cognitive-behavioural therapy for young children with anxiety disorders: Comparison of a Child + Parent condition versus a Parent Only condition. *Behaviour Research and Therapy*, *47*(8), 654–662. http://doi.org/10.1016/j.brat.2009.04.008

Whalen, D. J., Sylvester, C. M., & Luby, J. L. (2017). Depression and Anxiety in Preschoolers: A Review of the Past 7 Years. *Child and Adolescent Psychiatric Clinics of North America*, *26*(3), 503–522. http://doi.org/https://doi.org/10.1016/j.chc.2017.02.006

Wichstrøm, L., Belsky, J., & Berg-Nielsen, T. S. (2013). Preschool predictors of childhood anxiety disorders: A prospective community study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(12), 1327–1336. http://doi.org/10.1111/jcpp.12116

Part 2: Empirical paper

The Development of the “P-Factor” in Early Childhood: Examination of Genetic and Environmental Indicators.

Abstract

BACKGROUND: There has been increasing interest in the reconsideration of psychological disorders, moving away from traditional diagnostic approaches and towards models that reconceptualise psychopathology into a general psychopathology dimension ("p-factor") and additional specific dimensions. A key limitation of research in this area is that few studies have investigated how the p-factor develops over time during childhood, specifically during the preschool period.

AIM: To conduct the first longitudinal study of the evolution of the p-factor that begins in early life using a genetically sensitive design. In particular, this study seeks to examine the strength consistency of the p-factor in early childhood and examine genetic and environmental contributions to variability in the p-factor during this early stage of the lifecycle.

METHOD: Using longitudinal data from the Early Growth and Development Study, Confirmatory Factor models and Confirmatory Bi-factor models were applied to the CBCL from age eighteen months to seven years. Data were also obtained regarding birth mother psychiatric status using the CIDI to estimate genetic influences on child psychopathology symptoms. Adoptive parent reports of state anxiety and depression, marital hostility and economic distress were used to estimate environmental effects.

RESULTS: The p-factor appeared to increase in strength over time, with the variance accounted for by the specific factors declining, particularly after 27 months. Models of genetic influence of child p-factor nearly reached statistical significance at 27 months and at 4.5 years. Robust effects of the environment on child p-factor development were found across all ages.

CONCLUSION: The current findings suggest that the balance of general and specific factors in child psychopathology may change with development. There was weak evidence of genetic influences on p-factor over time, which could be a result of

methodological limitations. Additionally, the impact of adoptive maternal psychopathology was found to significantly contribute to child p-factor.

Introduction

The importance of understanding the early development of psychopathology.

The presence of behaviour problems in early development have been linked to later psychopathology and a wide range problematic outcomes in later life. Early behaviour problems tend to be clustered into two broad factor dimensions: externalising problems (which may feature disruptive and/or impulsive behaviour and can include substance use) and internalising problems (which features a large range of behaviours and disorders including somatic, depressive and anxious symptoms)(American Psychiatric Association, 2013). Prevalence rates indicate that such behaviour problems are common from preschool (Black, Jukes, & Willoughby, 2017) and continue to be prevalent in childhood; with estimates for "externalising" disorders like attention-deficit/hyperactivity were 8.6% and 2.1% for conduct disorders, for "internalising" disorders estimates were 3.7% for mood disorders, and 0.7% for panic disorder or generalized anxiety (Merikangas et al., 2010).

Whilst it is commonly expected that some behaviour difficulties in preschool years are a normative part of development, the evidence shows that for some toddlers and pre-schoolers they will continue to experience further difficulties throughout their development. Such studies have shown persistence of these disorders over time, with lifetime prevalence rates ranging from approximately 15%-25% for "externalising" disorders and 21-29% for "internalising" disorders (Kessler et al., 2005). By the age of 14 years 50% of 'lifetime' cases will be diagnosed; increasing to 75% by the age of 24 (Kessler et al., 2005). The stability of these behaviour problems highlights a clear need for a better understanding of the potential mechanisms responsible for the emergence and persistence over time.

Additionally, the impact of behaviour problems at an early age has far-reaching consequences on the child's functioning in many different domains. For example, high levels of physical aggression in toddlers aged 3-4 years had been

found to predict social and academic functioning at age 12 (Campbell, Spieker, Burchinal, & Poe, 2006). Such knock-on consequences, or developmental cascades, are not transient and could account for the long-lasting impact of childhood problems on widespread difficulties in development (Masten & Cicchetti, 2010). Developmental cascade models for antisocial behaviour have been well supported by developmental literature; such as the "dual failure" theory that hypothesises that early externalising behaviours continue in school, resulting in failure of social and academic functioning, that causes internalising problems and further antisocial behaviour as a child becomes more associated with deviant peers (Capaldi & Stoolmiller, 1999; Patterson & Stoolmiller, 1991). Thus it is of crucial importance to understand the aetiology of these early behaviour problems to help develop ways to intervene and prevent this negative trajectory.

What do we know about the genetics of childhood behavioural problems?

Developmental literature examining the potential mechanisms driving the development and persistence of behavioural problems has focused on the extent of the contributions of both genetic and environmental influences, which appears to differ by disorder type.

For externalising behaviours, heritability estimates for the clinical diagnosis of conduct disorder are moderate (Polderman et al., 2015; Salvatore & Dick, 2016) and comparable to the estimates of the related broader construct of antisocial behaviour, where meta-analyses of twin and adoption studies have found moderate heritability of around 40-50% (Moffitt, 2005; Rhee & Waldman, 2002). The heritability of substance use, depending on criteria, can range from 25 to 50%; with genetic effects being much weaker for initial experimentation and higher for substance abuse and dependence (Rutter, 2006). Adoption studies have found two genetic pathways to substance abuse: one indirectly via adoptee aggressiveness

and antisocial behaviour, and the other directly related to the biological parent's alcoholism (Cadore, Yates, Ed, Woodworth, & Stewart, 1995).

For internalising disorders, heritability estimates are more varied. For depressive disorder heritability has been estimated at around 40%, although this is dependent on both the population studied and measurement used (Polderman et al., 2015; Rutter, 2006; Sullivan, Neale, & Kendler, 2000). Such studies indicate that, whilst depressive disorder is clearly influenced by genetics, overall the environment has more of an influence in population variation for depression. Similarly, estimates of heritability for anxiety disorders from meta-analyses of twin and family studies range between 25-43%, depending on the disorder (Hettema, Neale, & Kendler, 2001; Scaini, Belotti, & Ogliari, 2014; Shimada-Sugimoto, Otowa, & Hettema, 2015; Van Houtem et al., 2013), with genetic and non-shared environmental factors explaining the majority of the variance. Family studies show moderate aggregation of all anxiety disorders (Hettema et al., 2001; Shimada-Sugimoto et al., 2015).

Heritability estimates appear to change over time. Generally, there is a tendency for heritability to increase with age (Bergen, Gardner, & Kendler, 2007; Polderman et al., 2015), .A large meta-analysis, that collated data from early childhood to adulthood of twin studies from the past fifty years, found an increase in heritability for conduct disorder (h^2= 50% to 66%); depressive episode (h^2= 36% to 42%), anxiety disorders (h^2= 42% to 56%) and alcohol disorder (h^2=42% to 50%)(Polderman et al., 2015). These findings were consistent with an earlier meta-analysis examining the cross-time heritability of disorders from late adolescence to adulthood. This study found that externalising behaviours showed a moderate significant increase in heritability over time; for internalising behaviours, the largest increase in heritability over time was found for anxiety, with depressive symptoms demonstrating a moderate but significant increase per year (Bergen et al., 2007).

Few studies have examined genetic influence very early in development. Studies that have, have demonstrated some genetic influences for behaviour problems for 2-3 year olds, accounting for on average 64% of the variance (van den Oord, Verhulst, & Boomsma, 1996). Research into the heritability of different anxiety phenotypes (e.g. inhibited, obsessive-compulsive, etc.) in preschool children has found the influence of genetics to vary considerably (h^2 =39% to 64%) and with high genetic overlap between general distress and the other anxiety phenotypes (Eley et al., 2003). Additionally, genetic factors have been found to underlie the frequency and stability of physical aggression during early childhood, with limited influence of the environment (Hudziak et al., 2003; Lacourse et al., 2014)**.** Thus, it appears that there are initially high heritability rates for psychopathology in the preschool years, which appears to then decrease in childhood then increase again over later life. However, the limited extent of evidence in the very early years limits how strongly this conclusion can be made.

Overall, the heritability findings quite clearly indicate an individual's liability to develop externalising and internalising disorders are both influenced by genetic and environmental factors, with the magnitude of genetic influence varying considerably over time. These high rates of heritability of behaviour problems, indicate the important influence of genetics, though relatively little is known about the role of genetics in very early development.

Challenges of using the current nosology to understand development of psychopathology

In recent years, extensive evidence has accumulated which indicates the substantial overlap between, or co-occurrence of many emotional and behavioural problems, and highlighted the potential importance of common biological factors underlying comorbidity of psychopathology (Kendler, Prescott, Myers, & Neale, 2003; Lahey, Van Hulle, Singh, Waldman, & Rathouz, 2011). The high frequency of

comorbid mental health disorders is well known (Hasin & Kilcoyne, 2012), to the extent that almost half of individuals diagnosed with one disorder will meet the criteria for another (Caspi, Houts, Belsky, & Goldman-mellor, 2015; Newman, Moffitt, Caspi, & Silva, 1998).

Certainly it has been a challenge for the field of molecular genetics to identify singular DNA risk variants (single-nucleotide polymorphisms; SNPs) for specific psychiatric disorders, with genome-wide studies demonstrating instead the highly polygenic nature of disorders (with numerous combinations of common variants of SNPs, mainly each individually with a small effect) and cross-disorder genetic overlap for a wide range of disorders (Smoller et al., 2018). In other words, the liability to develop a psychiatric disorder, as presently defined by psychiatric nosology systems (e.g. DSM), is influenced by multiple genetic risk factors that also share some genetic risk for the development of other psychiatric disorders.

This pervasive cross-disorder genetic risk highlights a major critique and problem of the psychiatric nosology systems. These high rates of comorbidity of externalising and internalising disorders are a concern for studies that are seeking to identify etiologic causes of specific disorders, because by ignoring the comorbidity, problematic assumptions are made of whether the disorder is the same regardless of the presence of a comorbid disorder, or whether the correlates being examined are related to a comorbid disorder (Caron & Rutter, 1991).

Indeed, longitudinal examination of the latent profiles of comorbid, internalizing and externalizing symptoms conducted across the first three years of school revealed a markedly high continuity for the comorbid symptom profile (89%), with 25% of an externalising-only profile transitioning to showing comorbidity and 20% of the internalising-only profile transitioning to a well-adjusted profile (Willner, Gatzke-Kopp, & Bray, 2016). This implies the potential for distinct pathways in the

development of behaviour problems, and a common vulnerability to developing comorbidity that is present and stable from early childhood.

Furthermore, research has indicated genetic factors that underlie comorbidity (Kendler et al., 2003; Lahey et al., 2011; O'Connor, Neiderhiser, Reiss, Hetherington, & Plomin, 1998). A population based twin study found two common genetic factors for comorbidity: one factor loading onto all three internalising disorders (generalised anxiety disorder, phobia and major depression) and another factor loading onto all four externalising disorders (drug abuse/dependence, alcohol dependence, conduct disorder and adult antisocial behaviour); demonstrating a large amount of shared heritable risks underlying comorbidity within these two broad band internalising and externalising subgroups (Kendler et al., 2003).

Yet there is also substantial comorbidity between internalising and externalising problems, with the two factors being highly correlated (~.5) (Krueger & Markon, 2006). What is more, common genetic factors have been implicated in the comorbidity across internalising and externalising factors; showing even at a genetic level that some disorders are highly related, with genetic correlations of approximately .7 (Smoller et al., 2018). Furthermore, longitudinal examination of the co-occurrence of the externalising and internalising dimensions in adolescent youth has also found genetic factors account for most of the stability and co-occurrence of depression and antisocial behaviour (O'Connor et al., 1998). Finally, a large family study using data from the Swedish National Register of those diagnosed or treated for psychiatric disorders found a common genetic origin across all disorders and criminal activity (Pettersson, Larsson, & Lichtenstein, 2015).

To summarise, given the extensive comorbidity and shared genetic effects, there is increasing evidence that individual disorders, and even broad-band internalizing and externalizing disorders, may not be the best way of conceptualising the structure of psychopathology in adulthood or in childhood. If this is the case,

findings from previous genetic studies on etiologic factors in the development of behaviour problems are limited because of the poor operationalisation of the relevant traits.

The p-factor as an alternative structure of psychopathology

Responding to these findings, researchers in the last decade have sought to develop a more parsimonious model of psychopathology. This has resulted in the proposal of an alternative structure of psychopathology, whereby in addition to specific factors underlying psychopathology (e.g. propensity to develop internalising or externalising disorders) there is a general vulnerability to psychopathology factor (later known as the "p-factor") orthogonally related to the specific factors (Caspi et al., 2014; Lahey et al., 2012). When compared to the correlated factor models, the bifactor model (which contains the additional "p-factor" with all disorders loading directly onto it) consistently fits the data better than a more traditional 'internalizing versus externalizing' model (Caspi et al., 2014; Lahey et al., 2012).

As many studies have demonstrated, the relationship between the specific factors (notably externalising and internalising) changes with the addition of the p-factor, with the correlations becoming weaker and even reversing in sign (Caspi et al., 2014; Neumann et al., 2016; Waldman et al., 2016;Tackett et al., 2013). This implies that the positive correlation that is found between externalising and internalising problems could exist because of a more general liability to psychopathology (p-factor); once this risk is controlled for, it appears that those prone to internalising disorders (e.g. depression) are less prone to developing externalising disorders (e.g. conduct disorder), and vice versa (Caspi et al., 2014).

Since the discovery of a general psychopathology dimension (Caspi et al., 2014; Lahey et al., 2012), the p-factor has been well replicated across many different populations of adults (Lahey et al., 2012) and children (Martel et al., 2017; Neumann et al., 2016; Patalay et al., 2015; Waldman et al., 2016), including

longitudinal samples (McElroy, Belsky, Carragher, Fearon, & Patalay, 2017; Murray, Eisner, & Ribeaud, 2016). Across these studies the specific factors tend to consist of internalising and externalising indicators, with the addition of a third factor varying depending on the ranges examined, e.g. thought disorders (Caspi et al., 2014) or attention problems (Mcelroy et al., 2017). Although in some models alternative specific factors have fit the data best (e.g. Fear, Distress and Externalising; Martel et al., 2017).

Whilst evidence has illustrated the replicability of the p-factor, the question of what the p-factor represents still remains. It was originally proposed that it represents a tendency to experience co-morbid and persistent symptoms of psychopathology, akin to the notion of a general intelligence factor (Caspi et al., 2014). Some researchers have proposed that the p-factor has an underlying more substantive meaning, reflecting an impulsive responsiveness to emotion (Carver, Johnson, & Timpano, 2017).

Other researchers have criticised the p-factor as being a statistical artefact as a result of positive manifold in the data (positive correlation of all variables with each other), and have emphasised the need of research to find a referent of the p-factor that is external to the statistical model, in order to legitimise its existence (Bonifay, Lane, & Reise, 2017; van Bork, Epskamp, Rhemtulla, Borsboom, & van der Maas, 2017). This is particularly important for the interpretation of the p-factor and establishment of its clinical utility.

To date, the literature has found the p-factor to be strongly correlated with family psychiatric history (Caspi et al., 2014; Martel et al., 2017) and childhood maltreatment (Caspi et al., 2014); lower IQ (Caspi et al., 2014); more life impairment (Newman et al., 1998); higher levels of socioeconomic deprivation (Patalay et al., 2015); and lower executive function ability (Martel et al., 2017). The p-factor has also been shown to improve prognostic predictions of later psychopathology (Lahey

et al., 2012; Patalay et al., 2015) and future educational attainment (Patalay et al., 2015).

Whether the p-factor is merely statistical variance or represents a substantive construct continues to be deliberated in the literature. Nonetheless, the p-factor continues to be well-replicated, shows stability over time, has robust external correlates, and provides a good model for predicting future psychopathology.

Development of the P-factor

A key question for developmental research is how the p-factor develops over time. There have been contrasting processes that have been hypothesised. The first is known as "dynamic mutualism", which suggests that the p-factor develops as a result of the individual symptoms interacting locally with each other and developing an emergent network structure (whereby influence increases and becomes reinforced as symptoms interact with other "distinct" symptoms). In this model then the p-factor reflects the emergence of comorbidity, or co-occurrence, due to increasing associations and interactions between individual symptom domains (Murray et al., 2016; van der Maas et al., 2006). The second model is known as the "*p* differentiation" model, whereby the p-factor is thought to reflect a more general and broad risk of developing any psychopathology, with manifestations of particular presentations of psychopathology being a result of expressions becoming increasingly specific over time (Caspi et al., 2014; Murray et al., 2016).

There has been little research addressing this; so far only two early longitudinal studies of the phenotypic stability of the p-factor have been undertaken (Mcelroy et al., 2017; Murray et al., 2016). Murray and colleagues (2016) examined the stability of the p-factor using eight time points from the ages of seven to fifteen and, in contrast to the hypotheses outlined above, they found that the strength of the p-factor remained constant over this time period. This was consistent with other

studies that found strong continuity over a shorter time period (i.e. two assessment waves) (e.g. Snyder, Young, & Hankin, 2017). However, more recent work by McElroy and colleagues (2017) across a much wider age range and beginning very early in life (ages two to fourteen) found evidence for both p-differentiation and dynamic mutualism occurring over time. This latter study is the only study thus far to examine the development of the p-factor in the first five years of life.

In sum, to obtain a clearer understanding of the phenotypic stability and development of the p-factor, more research is needed with multiple assessment waves that include measurements that begin early in development.

Genetics of the p-factor

As outlined previously, there is strong evidence for the influence of genetics in the development of psychopathology. Most studies that have examined the heritability and influences of genetics have used the traditional structure of psychopathology based on the current psychiatric nosology system.

Modelling of the maternal and paternal p-factor has shown significant associations between parent p-factor and child p-factor (Martel et al., 2017); yet this only represents indirect evidence of genetic effects. However, a recent twin study estimated moderate heritability for the p-factor in children aged 9-17 (Waldman et al., 2016).

Furthermore, a recent genomic twin study, using data from pre-adolescent children, showed significant SNP heritability of the p-factor (h^2=38%), implying common SNPs are partly responsible for the shared genetic effects on the p-factor (Neumann et al., 2016). Another genome-wide association study suggested that the p-factor accounted for variance in the associations between the common genetic variants linked to ADHD and other neurodevelopmental problems, externalising problems and to some degree internalising problems; thus suggesting that a

substantial proportion of the genetics associated with risk in developing ADHD belong to a broad non-specific genetic liability towards child psychopathology (Brikell et al., 2018)

Another twin study explored whether the variance explained by the p-factor overlapped, on both genetic and phenotypic levels, with the personality trait of negative emotionality; and found the strongest correlations were between negative emotionality and p-factor compared to other dispositions (e.g. pro-sociality and daring) and the other specific factors of internalising and externalising (Tackett et al., 2013). These findings have yet to be replicated and it is not yet known whether estimates are representative across the lifespan, particularly early in development.

Current study

As outlined above, little is known about the development of the p-factor over time, particularly in the early stages of development, and nothing is yet known regarding the genetic and environmental influences on the p-factor at this age. The present study will utilise a longitudinal adoption study design, using data collected from birth, to examine the stability of p-factor over time and the genetic and environmental influences on it across development up to age 7 years.

Adoptive study designs play an important role in elucidating the nature of the contribution of genetics in the development of psychopathology. Genes operate both directly, through “additive” effects of genes and environment (Rutter, 2006), and indirectly through the environment by gene-environment (rGE) correlations (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983). These types of genetic mediation of the environment are thought to occur in three forms: passive rGE, active rGE and evocative rGE (Plomin et al., 1977; Scarr & McCartney, 1983). Passive rGE refers to the process that parents pass on the ‘risk’ through providing both genes and environment. In contrast active rGE refers to the propensity of the child with the genes to seek out particular environments and provide them with

experiences that can exacerbate or alleviate risk. Finally, evocative rGE works in a similar way to active rGE, however instead of the child's genes influencing the selection of environments, it affects the responses of other.

Without genetically sensitive designs the extent to which behaviour is influenced by genes or environments cannot be quantified. Crucially, as well as allowing the estimation of genetic influences on child development, adoption designs also remove any inflation in the apparent effects of the environment caused by passive gene-environmental correlation.

Whilst the heritability of p-factor has been explored in previous studies of adults and adolescents (Martel et al., 2017; Neumann et al., 2016; Waldman et al., 2016), to our knowledge this is the first study to examine the genetic influences of the p-factor in early development, and the first to use an adoptive study design.

To test for genetic influences on the p-factor we used interviewer-based psychiatric diagnoses of birth mothers, combined using the same modelling principles used to capture general and specific factors in psychopathology in children. We also considered three environmental risk variables that have been shown independently to impact on the development of child psychopathology: financial distress (Conger et al., 2016; Patalay et al., 2015; Stover et al., 2012); parental marital hostility (Conger, Ge, Elder, Lorenz, & Simons, 1994; Conger et al., 2016; Stover et al., 2012); and parental internalising behaviours (Bagner, Pettit, Lewinsohn, Seeley, & Jaccard, 2013; Goodman et al., 2011). These are common environmental risk factors that have consistently been implicated in the development of externalising and internalising disorders in children.

This present study aims to address the following questions:

- Does the p-factor model provide a superior fit to child psychopathology data than the internalising and externalising two-factor model in late infancy and early childhood?
- Does the strength of p-factor grow over time (dynamic mutualism) or become more specific over time (p-differentiation)?
- What are the genetic contributions to the p-factor? Is this consistent over time?
- How do the environmental risks affect the development of the p-factor? Does this change over time?

Method

Participants

Data for this study came from the Early Growth and Development Study (Leve et al., 2007; Leve & Neiderhiser, 2013), a prospective longitudinal adoptive cohort study. This study was designed to examine the interplay between environmental processes (e.g. family, peer) that are mediated or moderated by genetic influences. Recruitment to the study occurred over two cohorts primarily through 45 adoptive agencies across 15 states in the USA (Leve & Neiderhiser, 2013; Leve et al., 2007). There were no restrictions to the adoptive family types included this study. To be recruited into the study the adoption had to be a domestic placement with a non-relative adoptive family and the baby placed within 3 months of birth. Additionally, both the adoptive and birth family had to understand English to an eighth-grade level. Babies with major medical conditions (e.g. extreme prematurity or extensive medical surgeries) were excluded. The mean age of baby at placement was 6.2 days (*SD* = 12.45).

A total of 561 triads (birth mother, adoptive mother and adoptive child) were included in this study. In comparison to the birth parents, the adoptive parents

recruited to this study were more likely to be white, come from a higher educational background with a higher income and be married (see Table 1). Ethical approval was obtained from the Institutional Review Boards of all participating organisations (George Washington University, The Pennsylvania State University, University of California, Davis, University of Minnesota, Oregon Social Learning Center).

Table 1. Demographic information collected on adoptive and birth parent families.

	Adoptive parent 1	Adoptive parent 2	Birth mother	Birth Father
Mean age at adoption (*SD*)	37.4 (5.6)	38.3 (5.8)	24.4 (6.0)	26.1 (7.8)
Race (% Caucasian)	91.8	90.4	70.1	69.9
Mean educational level (*SD*)[a]	5.9 (1.3)	5.6(1.5)	2.6(1.3)	2.7(1.3)
Married (%) [b]	91.1	91.1	30.6	51.4
Median annual household income	$100K+	$100K+	<$15K	$15K-$25K

[a] on a 7-point scale from 1(<high school degree) to 7 (graduate program);[b] Includes living together in a committed "marriage-like" relationship.

Measures

Child Psychopathology

The Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2000, 2001) was used as a measure of child psychopathology. The CBCL is a standardised and well validated parent-report measure, with good reliability (Cronbach's alpha reported for the DSM-scales ranged from .72 to.91) (Achenbach & Rescorla, 2001). The CBCL was completed by the adoptive mothers at ages: 18 months, 27 months, 4.5 years, 6 years and 7 years old. The items are rated on a three point scale as "not true", "somewhat/sometimes true" and "very true/often". From the ages 18 months to 4.5 years adoptive mothers completed the preschool version containing 99 symptoms of psychopathology (CBCL/1.5-5; Achenbach & Rescorla, 2000) and the school age version was administered at age 7 containing 113 psychopathology items (CBCL/6-18; Achenbach & Rescorla, 2001). At age 6 years half of the sample

were administered CBCL/1.5-5 years in error thus this wave was excluded from our analysis.

The CBCL items from the DSM Oriented Scales were used to create the internalising and externalising specific factors. For the preschool CBCL/1.5-5, the Internalising factor comprised of Affective Problems (10 items) and Anxiety Problems (10 items), with the externalising factor including only Oppositional Defiant Problems (6 items) (see Figure 1, *Appendix D*). The externalising factor for the school age CBCL/6-18 comprised of Oppositional Defiant Problems (5 items) and Conduct Problems (16 items). The internalising factor at this age included items from the Depressive Problems (12 items) and Anxiety Problems (9 items) scales (see Figure 2, *Appendix D*).

Due to low numbers of endorsement of psychopathology items in this study and to be consistent with common practice for item-level analysis of the CBCL, items were recoded into binary categories to indicate the presence and absence of symptom, (Achenbach & Rescorla, 2001). To assist the model estimation, items with variance below the threshold of 0.05 were removed from the analysis.

Genetic indicators: Birth parent psychopathology

To explore the genetic influences on child psychopathology development, a replication of the p-factor model from Caspi and colleagues (2014) original study was applied to the internalising and externalising dimensions of the birth parent measures of psychopathology. The thought disorder factor was not able to be modelled, as data on mania and schizophrenia was not collected in the EDGS study. The parent p-factor model (see Figure 3, *Appendix E*) used the Composite International Diagnostic Interview (CIDI; Kessler & Üstün, 2004) and the supplement Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981) as a measure of birth mother psychopathology, administered when

the adopted child was 18 months. These measures provide data on lifetime presence/absence of diagnosis.

The CIDI (Kessler & Üstün, 2004) is a standardised, comprehensive diagnostic interview tool designed to be used by non-clinical to diagnose mental health disorders from the fourth revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA ref) and the tenth version of International Classification of Diseases (ICD 10; WHO ref). The literature reports good reliability and validity of the interview(*k* =.45 to .63) (Andrews & Peters, 1998). Only a subset of the items were used in EDGS.

For this study birth mother psychopathology was modelled using the following diagnoses from the CIDI: Alcohol Abuse, Illicit Drug Abuse, Tobacco Dependence, Major Depressive Disorder, Generalized Anxiety Disorder, Specific Phobia and Social Phobia. For the EDGS study, CIDI responses were coded as: "yes", "no", "don't know" and "refused". For the purposes of this analysis people that "refused" or that received "don't know" were recoded as missing data, whilst the remaining items were recoded into binary categories to indicate the presence and absence of DSM diagnosis.

The DIS has reported fair reliability (*k*=.63 for Antisocial Personality Disorder) (Robins, Helzer, Ratcliff, & Seyfried, 1982). In this study responses to the DIS were coded as: "negative", "all criteria met, with no previous conduct disorder diagnosis", "all criteria met *including* a diagnosis of conduct disorder present". The original adult p-factor model (Caspi et al., 2014) included Conduct Disorder, Alcohol and Drug Dependence, in this study these items were substituted for Antisocial Personality Disorder, Alcohol Abuse and Drug Abuse.

Due to large amounts of data missing, birth father data was not used in this study (birth father *N*=188).

Environmental indicators

Adoptive parent psychopathology

State measures of anxiety and depression collected at each time point were calculated for each adoptive parent. Becks Anxiety Inventory (BAI; Beck, A. T., & Steer, 1990) was administered at 18 months to 36 months and the State Trait Anxiety Inventory for Adults (STAI; Spielberger, Gorsuch, & R.E., 1970) at 7 years old. The BAI is a 21-tem questionnaire that asks participants to rate symptoms of anxiety (e.g. “feeling hot”, “shakiness”) on a 4-point Likert scale from “not at all” to “severely (I could barely stand it)”. It has been reported to have good internal consistency (α=. 9) and test re-test reliability (*r*=.75) for psychiatric patients (Beck, Epstein, Brown, & Steer, 1988). Whilst the BAI has been shown to have good discriminant validity, it performs less well on tests of the convergent and construct validity compared to the STAI (Beck, Epstein, et al., 1988; Creamer, Foran, & Bell, 1995). The STAI has good internal consistency (α =.94), superior to the BAI (α = .88) and test-rest reliability (*r*=.68 *cf.* *r*=.62) in a non-clinical population (Creamer et al., 1995). The Beck Depression Inventory (BDI; Beck & Steer, 1987) was administered across all time points. The BDI is a well-validated and widely used measure of depression with good internal consistency (α = .87; Beck, Steer, & Garbin, 1988).

Both the correlations and Cronbach’s alpha between measures of state anxiety and depression for each time point were high for adoptive mothers (*r*= .53 to .63, *p*<.01; α = .69 to .77) and adoptive father’s (*r*=.59 to .68, p<.01; α =.74 to .8). Scores from each parent were standardized and averaged to create a composite psychopathology factor. The decision was made to keep them as separate pathology scores for each parent (as opposed to a combined score) due to low correlations between the two scores (e.g. at 18 months *r*=.15, *p*=.004 ; α = .25). At each time point a mean score was created of the standardized psychopathology

score up until that point (e.g. at 4.5 years adoptive mother psychopathology score would include a mean of data at 18 months, 27 months and 4.5 years).

Marital conflict

The 13-item marital hostility subscale from the Behavior Affect Rating Scale (Melby, Conger, Ge, & Warner, 1995) was used as a measure of conflict. The adoptive mother and father were asked to rate on a likert scale (1=always to 7 =never) the frequencies of certain behaviours in their interactions with their partner over the past year (e.g. “Listen carefully to your point of view”, “Insult or swear at you”, etc.). Total scores of hostility received from partner were calculated. Each partners score were significantly correlated across each time point (r= .45 to .55, p<.001*)* and thus were averaged to create a composite score of marital conflict. For each time point a cumulative mean was created of all marital conflict data collected until that point (e.g. age 27 months would include the mean of 18 months and 27 months).

Economic distress

The Making Ends Meet subscale was used as a measure of economic distress, taken from an interview about family demographics and financial situation (Conger, Ge, Elder, Lorenz, & Simons, 1994). For this subscale each parent answered questions about whether they have difficulty paying bills over the last 12 months and making ends meet. Responses were coded on a 5-pont scale (1=great difficulty, 5= no difficulty). In line with other research from the EDGS study (Stover et al., 2012) the scores from the adoptive mother and father were averaged to create a composite measure of family economic distress. Each time point represents the cumulative mean of economic distress scores at that time point. The measure had satisfactory internal consistency for adoptive mothers and fathers over time (α = .7 to .8).

Covariates

Openness of adoption and selective placement

Intrinsic to the assumptions made by adoption design studies is that there is a separation of effects from the biological parents (genetic effects) and the adopted parents (environment effects) (Rutter et al., 2008). Therefore a major challenge to assumptions of heritability comes from selective placements (whereby the adoptive agency may try to match a child to adoptive parents based on similarity to birth parent characteristics) and from the openness of adoption (through the amount of knowledge about birth parents and contact with birth family) which could inflate the genetic estimates from the model (Leve & Neiderhiser, 2013).

Explorations of the associations between birth and adoptive parent characteristics (which included 132 variables related to personality, financial needs, cognitive functioning ,etc., deemed not to be impacted by evocative effects) have been examined previously (Leve & Neiderhiser, 2013). They found no evidence of systematic bias, with only three associations meeting statistical significance.

The examination of the levels of adoption openness within the study showed significant variation and subsequently has been included in all EDGS papers (Leve & Neiderhiser, 2013). Despite this, only one paper has found a significant association involving adoption openness within the models being investigated (Leve et al., 2012). In this study the contact between the birth mother and child was controlled for by using a standardised mean of the 18-month adoptive mother, adoptive father and birth mother ratings regarding the extent to which they perceived that the adoption was open on a 7-point scale ranging from 1 (very closed) to 7 (very open).

Perinatal risk

Another important factor that may impact on the validity of the inferences made about the genetic effects are perinatal factors, such as pregnancy

complications, etc. To control for obstetric complications, a composite measure commonly used in EDGS studies was utilised. It consisted of a comprehensive assessment of perinatal complications from the birth mothers from data obtained by a pregnancy screener and a pregnancy calendar method developed for the study (originating from Caspi et al., 1996). It included data collected on maternal risk factors (e.g. depressive mood) and pregnancy complications (e.g. exposure to drugs), labour and delivery complications (e.g., prolonged labour) and neonatal complications (e.g. low birth weight, prematurity) (for more details see Marceau et al., 2013).

Analysis

The analysis was conducted in three stages as detailed below. All Confirmatory Factor Analysis (CFA) and Confirmatory Bi-factor Models (CBM) were specified and estimated using Mplus 8 (Muthén & Muthén, 2017).These models used the robust weighted least squares estimator (WLSMV), a pairwise present approach deemed effective at managing data missing at random and considered to be the best estimator for categorical data modelling (Brown, 2014).

The goodness of fit statistics and factor loadings were used to assess the quality of the model fit; this included the comparative fit index (CFI), Tucker Lewis Index (TLI) and root mean error square of approximation (RMSEA). Scores closer to 1 on the CFI and TLI indicate a good fit to the data, whereas a score of closer to 0 indicates a good fit for the RMSEA. To be considered a very good model fit for this study, the conservative and widely used criteria of CFI and TLI > .95 and RMSEA of <.06 were applied (Hu & Bentler, 1999).

Previous studies investigating the p-factor have found an effect of gender, specifically in becoming more pronounced in relation to the specific factors of internalising and externalising dimensions once the p-factor is accounted for

(whereby the magnitude of the correlation increased for both externalising and males, and internalising and females)(Caspi et al., 2014; Patalay et al., 2015). Consequently, gender was controlled for in all analyses of genetic and environmental influence.

Step one: The structure of child psychopathology

First, to examine the structure of child psychopathology and the fit of the p-factor model to the data, CFA and CBM were conducted at each time point. In line with the aforementioned literature (e.g. Caspi et al., 2014; Patalay et al., 2015), the CBMs contained a general factor ("p-factor") uncorrelated to the specific factors, with the correlations between specific factors and p-factor fixed to zero (see Figure 4). The specific factors in this study were Internalising and Externalising, no third factor was specified, to be consistent with the birth parent p-factor modelling in step two. The child psychopathology models used item level data (see Figure 1 and 2, *Appendix D*), while the parent psychopathology analyses relied on diagnosis-level scores. Explained Common Variance (ECV) was extracted for each factor to examine the strength of both the general p-factor and the specific factors over time. The ECV was originally developed to test the unidimensionality of a psychometric scale (Reise, Moore, & Haviland, 2010) and is calculated by dividing the variance explained by the general factor with the specific factors and general factor combined.

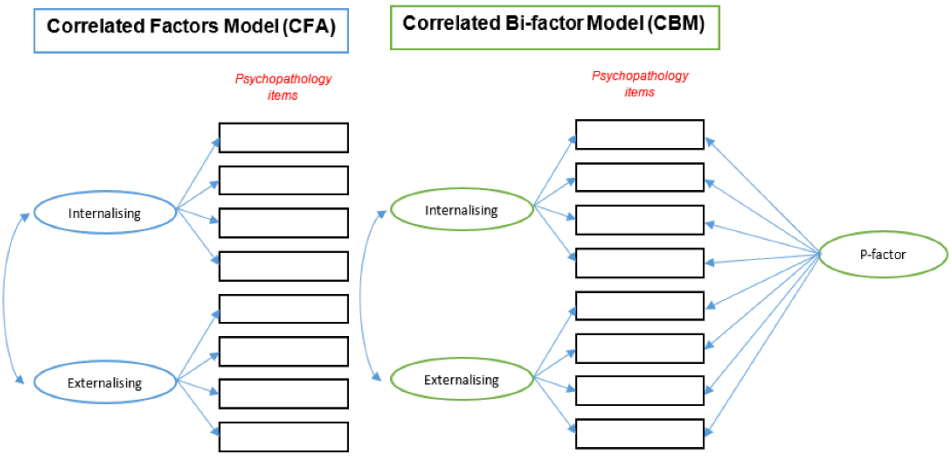


Figure 4. Diagram illustrating the structure of the correlated two-factor model and bi-factor models applied to child and birth parent psychopathology data.

Step two: modelling birth parent psychopathology

To create robust measures of birth parent psychopathology (i.e., the index of genetic risk), both the correlated and bi-factor models were applied to birth mother psychopathology data (see Figure 3, *Appendix E*). The goodness of fit statistics and ECV were examined and compared for both models. The resulting p-factor and specific factor scores were retained and used as measures of genetic influence.

Step three: testing genetic and environmental associations between birth parent p-factor and adopted child p-factor

Finally tests of environmental and genetic influence were examined. The aforementioned genetic and environmental variables were inputted into separate hierarchical regressions with the child's global psychopathology score (p-factor score) for each time point as separate dependent variables. Other covariates considered as potential confounds (e.g. openness of adoption, gender, etc.) were controlled for in all the regression analyses.

The hierarchal regression analyses were conducted in SPSS (Version 24). Data was checked for outliers and the analyses reproduced by removing extreme outliers (z=+/-3.29) using winsorizing (replacing value with the nearest reported value below the threshold) which produced no change to the interpretation of the

findings. Additionally, to test the robustness of the analyses to deviations from normality, the regression analyses were re-ran using a bias corrected accelerated (BCa) confidence interval set at the 95% percentile on a 1000 bootstrap samples (Efron & Tibshirani, 1994). Generally the results were similar for the bootstrapped and parametric analyses.

Results

The structure of child psychopathology

Overall model fit

The first CFA model applied to the child psychopathology data was the correlated two-factor model with two latent variables representing dimensions of internalising and externalising problems. The correlated model did not provide a good fit to the data at 18 months; RMESA was 0.06 and the CFI (.86) and TFI (.84) were below the threshold for acceptable model fit. The model partially met criteria for good fit at 27months, with acceptable RMSEA of .05, yet remained a poor fit according to the CFI (.88) and TLI (.87). At no point did the correlated factors model meet the criteria set for very good model fit (see Table 2). The best fitting model was at 7 years, where the CFI (.94), TLI (.93) and RMSEA (.04) revealed an overall acceptable model fit.

Compared to the correlated model, the bi-factor model provided a better fit to the data at each time point (see Table 2), with the later ages meeting the criteria for very good model fit, CFI (.96), TLI (.95) and RMESA at .03 for ages 4.5 and 7 years old.

Table 2. Table showing summary of goodness of fit statistics of CFA and CBM applied to child psychopathology over time

Age	*Model*	*X^2*	*df*	*CFI*	*TLI*	*RMSEA*
18 months	Correlated	394.89*	134	.86	.84	.06
	Bifactor	226.89*	116	.94	.92	.04
27 months	Correlated	414.62*	169	.88	.87	.05
	Bifactor	278.26*	149	.94	.92	.04
4.5 years	Correlated	505.54*	274	.91	.90	.05
	Bifactor	358.01*	249	.96	.95	.03
7 years	Correlated	596.76*	376	.94	.93	.04
	Bifactor	485.78*	347	.96	.95	.03

**p<.001*

Variance explained by the bi-factor model.

Both the variance and statistical significance for the p-factor varied over time, with the p-factor not reaching statistical significance at 27 months (*p*=.11). In general, the p-factor appeared to increase in strength between 18 months and 7 years, with the variance accounted for by the specific factors declining, particularly after 27 months (see Figure 5).

In comparison, when the correlated two-factor model was applied to the child psychopathology data, the proportion of ECV variance for the internalising and externalising factors remained consistent over time, with externalising explaining the most variance over time (see Table 3, *Appendix F*).

Notably, in the p-factor models the internalising factor did not reach statistical significance at any time point (see Table 5, *Appendix F*). The externalising factor, on the other hand, remained significant over time (*p*<.05-.001), accounting for a large amount of variance over time (σ^2 =.48 - .56) until 7 years when it no longer met statistical significance and the p-factor accounted for the majority of the variance (see Figure 5).

It was interesting to note that the individual items that loaded most heavily on the p-factor were items related to negative and angry affect (e.g. “enjoys little”, “cries”, “temper” and “looks unhappy”), particularly in the later ages (see Tables 6 to 9, *Appendix F* for CBCL item loadings).

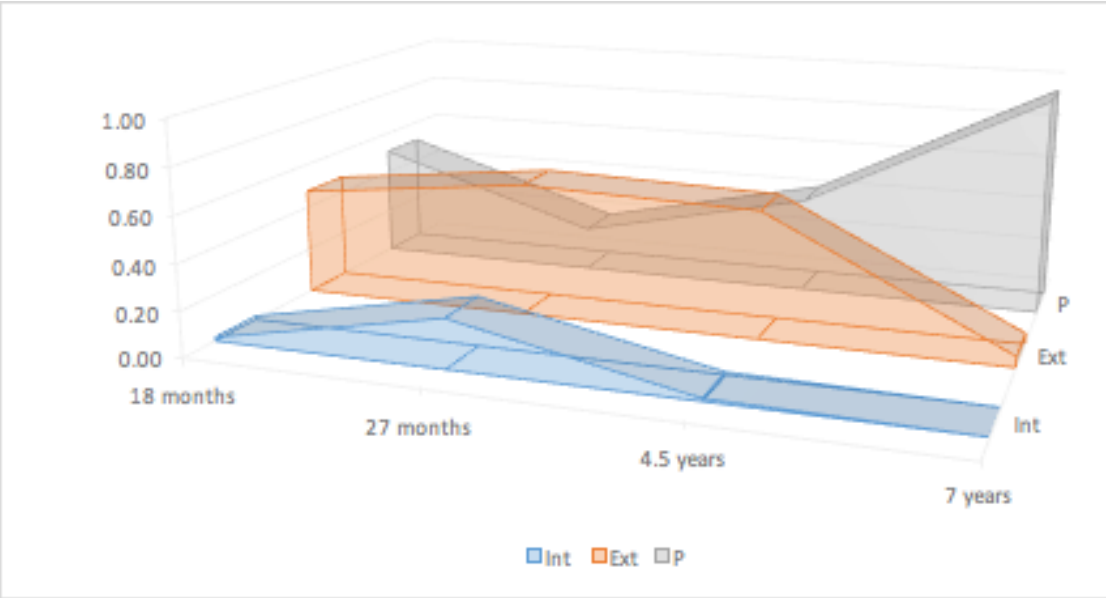


Figure 5. Graph showing Explained Common Variance (ECV) for general p-factor and specific factors of internalising and externalising over time.

Genetic and environmental contributions

Genetic influences I: modelling birth parent psychopathology

Before investigating the genetic contributions of the p-factor, both the correlated two-factor (internalising and externalising) model and bi-factor model were estimated from the birth mother psychopathology data (Figure3, *Appendix E*).

The birth mother bi-factor model was a good fit to the data ($X^2(11)=9.82$, $p=.55$; CFI=1, TLI=1.01, RMSEA=0) and was a superior fit to the data than the correlated factors model ($X^2(19)=40.53$, $p<.001$; CFI=.95, TLI=.93, RMSEA=.05). Like the child data, in the p-factor model of birth parent psychopathology, the internalising factor showed no significant variance ($\sigma^2 = .08$, $p=.55$), while the

externalising (σ^2=.27, *p*=.004) and p-factor (σ^2=.35, *p*=.01) accounted for the majority of the variance (*Appendix H*).

Genetic Influences II: testing associations between birth parent p-factor and adopted child p-factor

To examine the genetic contributions to child global psychopathology, a hierarchical regression was conducted, with all three factors from the birth mother bi-factor model predicting the child p-factor for each time point. The covariates were entered into the first block, followed by the birth mother internalising, externalising and p-factor scores entered together as one block to examine genetic influence of the p-factor.

The regression models investigating main effects of genetic influences on child p-factor did not show a consistent main effect of genetics over time, however ages 27 months and 7 years just missed statistical significance using the conventional 5% threshold (see Table 11). Birth mother p-factor was a positive significant predictor of child-p-factor at 27 months (*B*=.09, *p*=.04) and nearly met statistical significance at 4.5 years (*B*=.14*, p*=.06). These findings were similar to the bootstrapped analyses, which showed highly similar levels of statistical significance for both time points (*p*=.05). No main effects of birth mother externalising or internalising factors on child p-factor were found. At 7 years obstetric and perinatal risk factors became a significant positive predictor of child p-factor (*B*=.03, *p*=.03).

Table 11. Summary of the parametric and bootstrapped hierarchical regression models of genetic predictors for child p-factor over time.

	Parametric				*Bootstrap*			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
18 months (N= 470)									
Openness[1]	.03	.02	.07	.27	.03	.02	.16	-.01	.07
Perinatal risk[1]	.00	.02	-.01	.86	.00	.02	.86	-.03	.03
Gender[1]	.00	.04	.00	.68	.00	.04	.94	-.07	.08
BMint[2]	.01	.15	.00	.39	.01	.16	.93	-.30	.32
BMext[2]	.02	.07	.01	.38	.02	.07	.79	-.12	.16
BMp[2]	.08	.06	.07	.94	.08	.06	.18	-.02	.18
Δ R^2	.01								
F	.87			.46					
27 months (N=447)									
Openness[1]	.02	.02	.07	.14	.02	.02	.15	-.01	.06
Perinatal risk[1]	0	.01	0	.93	.00	.01	.94	-.02	.02
Gender[1]	0	.03	0	.99	.00	.03	.99	-.06	.07
BMint[2]	.07	.11	.03	.54	.07	.10	.50	-.11	.28
BMext[2]	.06	.05	.05	.31	.06	.05	.29	-.05	.16
BMp[2]	**.09**	**.04**	**.11**	**.04**	.09	.04	.05	.00	.17
Δ R^2	.02								
F	2.69			.05					
4.5 years (N=379)									
Openness[1]	.03	.03	.06	.27	.03	.03	.26	-.02	.08
Perinatal risk[1]	0	.02	.01	.85	.00	.02	.87	-.03	.05
Gender[1]	-.01	.05	-.01	.88	-.01	.05	.88	-.10	.09
BMint[2]	-.34	.19	-.10	.08	-.34	.19	.08	-.70	.04
BMext[2]	-.07	.09	-.04	.43	-.07	.09	.42	-.26	.10
BMp[2]	.14	.07	.11	.06	.14	.07	.05	.00	.29
Δ R^2	.01								
F	1.75			.16					
7 years (N=374)									

Openness[1]	.01	.02	.02	.7	.01	.02	.67	-.04	.05
Perinatal risk[1]	**.03**	**.01**	**.11**	**.03**	**.03**	**.01**	**.03**	.00	.06
Gender[1]	-.04	.04	-.05	.31	-.04	.04	.32	-.11	.03
BMint[2]	.28	.15	.10	.07	.28	.15	.07	-.04	.58
BMext[2]	.04	.07	.03	.58	.04	.07	.55	-.09	.17
BMp[2]	.07	.05	.07	.21	.07	.05	.19	-.05	.17
Δ R^2	.02								
F	2.5			.06					

SE*B*= standard error *B;CI*=confidence interval; [1]=entered in block one of hierarchical regression; [2]=entered in block two. Bold font indicates significant at p<.05.

Environmental influences on adopted child p-factor

Similar to examining the genetic indicators, the environmental predictors were also entered into a hierarchical regression after controlling for the covariate variables. Main effects of environmental influence on the child p-factor were found at each time point, with regression models illustrating strong statistical significance across each age ($p \leq .001$).

Across all ages adoptive mother psychopathology was a significant positive predictor of the p-factor (see Table 12). At age 27 months adoptive father psychopathology positively predicted child p-factor (*B*=.04, *p*=.04); and remained close to statistical significance in the bootstrapped sample (p=.06). In the older ages there was a positive trend of marital hostility predicating child p-factor for 4.5 years and 7 years that was just below statistical significance (*B*=.01, *p*=.06 and *B*=0, *p*=.06 respectively); although this trend was not found in the bootstrap analysis. Again, at 7 years obstetric and perinatal risk factors remained a significant predictor of child p-factor (*B*=.03, *p*=.01).

To check for evocative rGE influencing the environment measures, Pearson's correlations were conducted between birth mother p-factor and specific factors (internalising and externalising) and the environmental variables (see Table

13, *Appendix I*). With the exception of birth mother externalising factor predicting adoptive mother psychopathology at the 7 year time point, no other correlations were statistically significant.

Table 12. Summary of the parametric and bootstrapped hierarchical regression models of environmental predictors for the child p-factor over time.

	Parametric				*Bootstrap*			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
				18 months (N= 461)					
Openness[1]	.02	.02	.04	.39	.02	.02	.41	-.03	.06
Perinatal risk[1]	0	.01	.01	.89	0	.01	.88	-.02	.03
Gender[1]	0	.04	0	.91	0	.04	.92	-.09	.07
AM psych[2]	**.14**	**.02**	**.27**	**<.001**	**.14**	**.02**	**<.001**	.09	.19
AF psych [2]	-.03	.02	-.06	.21	-.03	.02	.22	-.07	.02
Economic distress[2]	.02	.02	.07	.14	.02	.02	.14	-.01	.05
Marital Hostilty[2]	0	0	-.02	.73	0	0	.71	-.01	0
Δ R^2	**10.29**								
F	**.08**			**<.001**					
				27 months (N=459)					
Openness[1]	.01	.02	.04	.42	.01	.02	.47	-.02	.05
Perinatal risk[1]	0	.01	.02	.66	.00	.01	.67	-.02	.03
Gender[1]	.01	.03	.01	.83	.01	.03	.85	-.06	.07
AM psych[2]	**.06**	**.02**	**.16**	**<.001**	**.06**	**.02**	**.01**	.02	.10
AF psych [2]	**.04**	**.02**	**.10**	**.04**	.04	.02	.06	0	.08
Economic distress[2]	.01	.01	.02	.61	.01	.01	.60	-.02	.04
Marital Hostilty[2]	0	0	-.06	.22	0	0	.2	-.01	0
Δ R^2	**.04**								
F	**4.56**			**.001**					
				4.5 years (N=394)					
Openness[1]	.03	.03	.06	.22	.03	.03	.20	-.02	.09
Perinatal risk[1]	0	.02	-.01	.86	0	.02	.87	-.04	.03
Gender[1]	.01	.05	.01	.81	.01	.05	.82	-.09	.1
AM psych [2]	**.11**	**.03**	**.17**	**<.001**	**.11**	**.04**	**<.001**	.04	.18
AF psych [2]	.05	.03	.07	.17	.05	.03	.17	-.02	.11
Economic distress[2]	-.02	.02	-.04	.49	-.02	.02	.52	-.06	.03

Marital Hostilty[2]	.01	0	.1	.06	.01	0	.08	0	.01
Δ R^2	**.06**								
F	**5.85**			**<.001**					
				7 years (N=392)					
Openness[1]	.03	.02	.07	.15	.03	.02	.14	-.01	.07
Perinatal risk[1]	**.03**	**.01**	**.12**	**.01**	**.03**	**.01**	**.02**	.01	.06
Gender[1]	-.02	.04	-.02	.63	-.02	.04	.63	-.09	.05
AM psych[2]	**.08**	**.03**	**.17**	**<.001**	**.08**	**.03**	**<.001**	.03	.14
AF psych [2]	.04	.03	.09	.11	.04	.02	.07	0	.09
Economic distress[2]	.02	.02	.06	.24	.02	.02	.27	-.02	.06
Marital Hostilty[2]	0	0	.03	.58	0	0	.58	0	.01
Δ R^2	**.06**								
F	**6.22**			**<.001**					

AM psych= adoptive mother psychopathology; AF psych=adoptive father psychopathology; SE*B*= standard error *B*; [1]=entered in block one of hierarchical regression;[2]=entered in block two. Significant results are highlighted in bold.

Secondary analyses: examining the genetic and environmental influences on the specific internalising and externalising factors of the bi-factor model

The regression models revealed no evidence of genetic influence on the specific factors of child internalising or externalising problems, with no model reaching statistical significance at any age. It was notable, however, that female gender negatively predicted child externalising at 7 years (*B*=-.02, *p*<.001) (Tables 14 & 15, *Appendix J*).

In comparison, the models of environmental influence provided better predictors of the specific child internalising and externalising factors, meeting statistical significance at ages 27 months and again at 7 years for child internalising only (see Tables 16 & 17 *Appendix K*). Similar to the findings of the child p-factor (although not as consistent) adoptive mother psychopathology was the most frequent predictor of internalising and externalising factors, positively predicting internalising factor at 27 months (*B*=.09, *p*<.001) and 7 years (*B*=.01, *p*=.003) and the externalising factor at 18 months (*B*=.04, *p*=.02), 27 months (*B*=.17, *p*<.001) and

4.5 years (*B*=.08, *p*=.01). Marital hostility was found to positively predict child internalising at ages 18 and 27 months (*B*=0, *p*=.03 and *B*=0, *p*=.03 respectively). Some of the covariates were also found to be significant predictors of child internalising and externalising. Similar to the genetic regression models, female gender was also found to be a negative predictor of externalising at age 7 (*B*=-.02, *p*=.002). Additionally, openness of adoption was found to be a positive predictor of externalising behaviour at age 4.5 years (*B*=.05, *p*=.04), yet a negative predictor of internalising at age 7 (*B*=-.01, *p*=.02). These results did not change for both genetic and environmental regression models in the bootstrapping analysis.

Discussion

Development of the p-factor

This current study sought to examine the structure and stability of psychopathology early in development and clarify the contributions of genetic and environmental influences in this process. This paper aimed to replicate the p-factor model (Caspi et al., 2014; Lahey et al., 2012) that has been well-replicated in cross-sectional data in older children and adults (Martel et al., 2017; Patalay et al., 2015; Waldman et al., 2016) and add to the literature examining the stability of p-factor over time (Mcelroy et al., 2017; Murray et al., 2016). In line with the aforementioned literature, compared to the correlated internalising and externalising two-factor model, the p-factor model provided a superior fit to the child psychopathology data from ages 18 months to 7 years.

Previous research examining the developmental stability of the p-factor found both the p-factor and specific factors remained consistent in the proportions of variance explained over time (Mcelroy et al., 2017; Murray et al., 2016); thus providing a challenge to the hypothesised processes, dynamic mutualism and p-differentiation, to explain the development of the p-factor. In contrast, the present

study found that the magnitude of the p-factor varied across time (.20 to .95). Overall, there was a trend indicating growth of the p-factor over time, which would be more consistent with the hypothesised dynamic maturation of p-factor development.

It is important to highlight that the externalising factor also accounted for a non-trivial and stable amount of variance (.48 to .56) until age seven. This is consistent with earlier research that highlighted a distinct externalising symptom profile in kindergarten children that was moderately stable through to second grade, whereby there was a 25% probability of externalising profile children changing to a comorbid internalising externalising profile (Willner et al., 2016), a pattern predicted by cascade models of the development of psychopathology (Masten & Cicchetti, 2010) and the "dual failure theory" of antisocial behaviour (Capaldi & Stoolmiller, 1999; Patterson & Stoolmiller, 1991).

Additionally, it is important to note that the p-factor reduced in magnitude, from .51 at 18 months to .20 at age 27 months at which point it was no longer statistically significant. This could partly support the p-differentiation hypothesis, whereby the p-factor represents a propensity to develop specific forms of psychopathology (Murray et al., 2016), though the later time points saw a decline in the specific factors . Thus, the results of this study may suggest evidence that both p-differentiation and dynamic mutualism are in operation at different stages of development, or that there are concurrent developmental processes occurring over time (i.e. for some individuals their symptoms may differentiate over time and become more specific; whereas for others they experience more comorbidity over time)(McElroy et al., 2017).

Further elucidation of the exact processes are beyond the scope of this study as a consequence of the chosen methodology; specifically it is important to highlight that this present study utilises a cross-sectional p-factor at each time point,

consistent with previous research (Mcelroy et al., 2017; Murray et al., 2016). Therefore, the p-factor derived at each particular age may be age-specific; in other words, its meaning may vary across different time points. That said, an examination of the p-factor item loadings for this study indicated that the p-factor predominately related to negative and angry affect, particularly in the later ages. Indeed, this could suggest a profile of irritability that has been linked to both externalising and internalising disorders (Leibenluft, 2017); and is consistent with psychological explanations of the p-factor that have suggested p-factor is related to impulsive reactivity to emotion (Carver et al., 2017) and tendency to experience distress and negative affect (Tackett et al., 2013).

Genetic contributions of p-factor

This study also sought to elucidate the contributions of genetic influences on the p-factor over time. There has been little research to date into the role of genetics in the p-factor, although twin studies have shown moderate heritability (Waldman et al., 2016) and common SNPs (Neumann et al., 2016) linked to the p-factor. To our knowledge, this is the first study to utilise the adoptive study design to explore the contribution of genetics of the p-factor, particularly in the younger age range. Additionally, this study applied the bi-factor model to birth mother data, to elucidate the contributions of birth mother p-factor and specific factors to child psychopathology development.

Overall there was relatively weak evidence of genetic influences on child p-factor and specific factors. Although this study found that birth mother p-factor significantly predicted child p-factor at ages 27 months and reached near significant levels at 4 years, none of the regression models predicting genetic influences for child p-factor reached statistical significance. While it is possible that these findings could imply genetic innovation (genetic effects appearing at distinct points in

development), this should be interpreted cautiously, as it is likely that the effects are not significantly different to each other across ages.

It is important to recognise that the hereditary effects in the current study were based on a phenotype of only the birth mothers data, and therefore only half of the potential genetics effects were modelled. Additionally, despite using a relatively large sample size for the analysis (*N*=374 -470), this study would not have been adequately powered to detect genetic effects with small effect sizes (<.2), which may have been further reduced through the separation of genetic risk into three separate factors through the birth mother p-factor modelling (see Figure 6, *Appendix L*).

Secondary analyses revealed no evidence of genetic influences on the specific factors of child internalising or externalising problems at any age. This was not in line with expectations from previous studies modelling p-factor heritability, that found genetic contributions across all three factors, with the externalising factor having the highest heritability, followed by p-factor and then the internalising factor (Waldman et al., 2016). Those findings were based on a larger sample of twins (*N*=1568 sets of twins), therefore it could be the case that this study was inadequately powered to detect genetic influences in this age range. However, they also used an older age range (9-17 years), thus the results are not be directly comparable to the younger age range used in this study.

Environmental risk on p-factor

Finally this study sought to examine the influence of environmental risks on the development of the p-factor. This study found that environmental risks were significantly associated with the development of p-factor across all time points; in particular, adoptive mother psychopathology was the most stable and prominent predictor of child p-factor variance in comparison to the other environmental influences measured. For the specific factors, adoptive mother psychopathology

also predicted internalising and externalising problems, although not as consistently or across all time points. This is in line with previous literature that has indicated an important role for maternal psychopathology in the development of child behaviour problems (Bagner, Pettit, Lewinsohn, Seeley, & Jaccard, 2013; Goodman et al., 2011).

Interestingly, in this study adoptive father internalising psychopathology only predicted child p-factor psychopathology at 27 months. Developmental research has tended to focus on the effects of maternal psychopathology, with less attention paid to the role of paternal psychopathology. Typically studies have shown that mothers spend more time in caregiving tasks and are more accessible to their infants than fathers (e.g. Laflamme, Pomerleau, & Malcuit, 2002), thus it could be assumed that there is more opportunity for depression and anxiety to impact on toddler interactions via the mother. Furthermore, it appears that fathers become more engaged in child play and rearing activities in toddlerhood when the child can interact more independently (Laflamme et al., 2002), which could mean that anxiety or depression, which may impair interactions with the child and influence the child's vulnerability to psychopathology, may become more apparent at this age. This would be in line with previous research from EDGS study group, that found that adoptive father depression at nine months contributed to child externalising problems at 27 months of age (Pemberton et al., 2010). It is also possible that father psychopathology may be indirectly predicting child psychopathology, but through indirect routes that are not examined by this current study. The same research conducted by the EDGS study group found adoptive father depression at nine months predicted toddler externalising symptoms at 27 months, by contributing to maternal depression at 27 months (Pemberton et al., 2010).

In addition to these primary hypothesised associations, we also observed that at age seven, when the child p-factor was its strongest in magnitude, perinatal

risk significantly predicted child p-factor. It is known that the effects of stress experienced in early life on emotional regulation and the associated brain regions, have been shown to be more resistant to recovery, even after the removal of the stressor (Pechtel & Pizzagalli, 2012). Most of the birth mothers in this sample had experienced some type of risk exposure (e.g. alcohol use, delivery complications) and it was more common for birth mothers to have experienced multiple risks compared to the general population of the USA (Marceau et al., 2016). The association we observed between perinatal risk and child general psychopathology is in line with other EDGS reports that have shown an evocative effect of perinatal risk exposure on child psychopathology (e.g. Pemberton et al., 2010).

Internalising and externalising factors

Secondary analyses showed no evidence of genetic influences for either of the specific child psychopathology factors (internalising and externalising) across all time points. This was only partly consistent with previous research using twin studies that had found highest heritability for externalising and low levels of heritability for internalising problems (Waldman et al., 2016). Environmental influences appeared to account for internalising and externalising, although not consistently over time. Of note, marital hostility predicted child internalising at ages 18 and 27 months. This is in line with other research that has found both mothers' and fathers' marital hostility were linked to child internalizing problems, connected via parent and child hostility (Low & Stocker, 2005).

Additionally, openness of adoption was found to be a positive predictor of externalising behaviour at four and a half years, yet a negative predictor of internalising at age seven. This could be explained by a number of developmental processes: a preschool child's understanding of the meaning of adoption is limited, mainly consisting of understanding the language of adoption (i.e. how to talk about adoption) (Brodzinsky, 2011). Therefore at four and five years children will still be

developing their ability to understand and communicate their feelings about their adoption, which may influence the interpretation of their behaviour by their parents as externalising. Once the child reaches middle school and their cognitive abilities develop, they begin to understand that their birth parents could have had other options which can undermine their sense of self (Brodzinsky, 2011), thus leading to more internalising symptoms .

Limitations

In line with the methodology of other studies that have explored the longitudinal research of the p-factor (Mcelroy et al., 2017; Murray et al., 2016), the p-factor at each specific time point was assessed cross-sectionally and therefore reflects a general statistical summary of comorbidity across internalising and externalising symptoms. Thus, this study examined the consistency of the p-factor and specific factors over time on a broad statistical level, and consequently does not inform about the symptom level continuity and persistence across development.

Secondly, decisions made when modelling the p-factor may have had the potential to affect the results. The decision to include items based on the DSM-oriented scales was made to ensure more similarity with the modelling of the birth parent p-factor (that was derived from internalising and externalising DSM diagnoses). Inevitably this meant using a more restricted number of items in the model, particularly externalising items in the preschool CBCL, which will have reduced variability and lowered the power to detect genetic effects.

Furthermore, the child p-factor was calculated from a symptom questionnaire, as opposed to clinical diagnoses based on interviews (like the birth mother-factor). There are some advantages to a symptom level approach: it can capture meaningful variation spanning across clinical and subclinical levels of psychopathology - the dichotomy of which is largely artificial (see Rutter, 2006) – but also avoids the issue of “artificial comorbidity” caused by different diagnoses

having the same symptoms in common (Caron & Rutter, 1991). However, there have been concerns raised about the challenges of using non-clinical selected samples to measure clinical traits (Murray et al., 2016; Murray, McKenzie, Kuenssberg, & O'Donnell, 2014; Reise & Waller, 2009).

Additionally, in this study only the adoptive mother's CBCL was used. It is a well-known that informant agreement has been a longstanding problem for identifying child psychopathology, with the lowest agreement between the child and other informant (Achenbach, McConaughy, & Howell, 1987). Whilst for the age range covered in this study it would have only been possible to obtain adoptive father scores, the exclusion of this is a limitation of the study. That said, the highest agreement between raters is generally between parents (r=.60) (Achenbach et al., 1987) and, as highlighted by parental research, it is more common that mothers take a more active role in caregiving from the earliest ages which one could assume means that they would best placed to identify symptoms of emotional and behaviour problems (Laflamme et al., 2002). Additionally, other research has shown that the p-factor could be reproduced from multiple informant sources of psychopathology and remained a superior fit to the data (Tackett et al., 2013).

Still, the use of only one informant could mean that there is increased risk of bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) and/or lower power. In particular, the environmental influence of adoptive mother psychopathology on the p-factor should be interpreted with caution, as they could have been influenced by common rater effects, i.e. an artificial inflation of the covariance between the adoptive mother predictor variables and the child psychopathology outcome variable is produced due to the same respondent providing the measurement (Podsakoff et al., 2003).

Key implications

Whilst there are now numerous studies that have replicated the p-factor at different ages, this study shows evidence of the replicability of the p-factor at the youngest age to date, thereby making an important contribution to the p-factor literature. The p-factor offers an alternative way to operationalise psychopathology that is not limited by arbitrary diagnostic thresholds and has been shown to provide a good predictive model of later psychopathology (Caspi et al., 2014; Patalay et al., 2015). The identification of the presence of the p-factor in very young children is important, given both the prevalence of behaviour problems in young children and the negative impact of early behaviour problems on the child's trajectory in development. The p-factor may provide a way of identifying and targeting children that may be at the most risk (highest p-factor). The dimensional approach offered by the p-factor model may suggest that research should be refocused on to more trans-diagnostic mechanisms and mechanism-focused interventions for psychopathology (e.g. Carver et al., 2017).

Few studies have examined the development of the p-factor over time, and in contrast to those other developmental studies (Mcelroy et al., 2017), this study demonstrates a change in magnitude of p-factor over time, with p-factor decreasing between 18 and 27 months and increasing in magnitude thereafter until age seven. This suggests that the p-factor is developmentally dynamic. If replicated, such findings may assist in prevention and intervention studies by furthering understanding of the developmental course of psychopathology and the impact of the timing of interventions on developmental outcome.

This is the first p-factor study to examine the genetic influences on the development of the p-factor utilising an adoptive study design. This present study showed some, albeit modest and inconsistent, evidence of genetic influences on the p-factor in young children. These differences over time could be interpreted as

evidence of genetic innovation during toddlerhood, but may also reflect random statistical variability. Toddlerhood is a time of considerable development and neurological changes; which could imply a high risk period for children with high genetic risk. The identification of specific risk alleles and understanding of the exact mechanisms of their interaction with the environment, that operate to increase or reduce vulnerability, will further the current findings and assist in more prognostic understandings of pathology (Rutter, 2006).

Finally, the present study highlighted significant environmental contributions to the development of child psychopathology, most notably the impact of adoptive maternal psychopathology on the child p-factor. The association between maternal depression and subsequent child emotional and behaviour problems is well established (Goodman et al., 2011; Beck, 1999); this study therefore follows others in highlighting the important impact of maternal depression on child development. There is increasing recognition that adoptive mothers experience post-adoption depression at similar post-partum rates to birth mothers (Mott, Schiller, Richards, O'Hara, & Stuart, 2011; Senecky et al., 2009), and this study further emphasises that this should be something to be recognised and prioritised for treatment by mental health professionals, especially to prevent deleterious outcomes for the child.

Future research

In this study individual items that related to negative and angry affect loaded most heavily on the p-factor, yet there was some variation in this at each age. Future research employing an individual symptom level exploration using statistical methodology, like network analysis, would provide valuable information that would complement the current findings, and examine specific causal pathways of both the p-factor and different symptom trajectories over development (see for example McElroy et al., in press).

The interpretation of the mechanisms of development of the p-factor in this study were based on interpretation of the strength consistency (i.e. ECV) of the p-factor over time. Future research exploring the phenotypic continuity explicitly, for example through the application of cross-lagged models (e.g. Mcelroy et al., 2017), may further elucidate developmental processes.

Finally, this study utilised a genetically sensitive design to examine the additive effects of genes on the development of the p-factor; yet there are many direct and indirect ways in which genes may influence the development of behaviour problems that were not examined, in particular through gene-environment interactions. It could be possible that genetic vulnerabilities exert their influence only in the presence of particular risk environments (Moffitt, 2005).

Conclusion

This present study demonstrated the p-factor to be a useful construct to examine the development of child psychopathology over time in very young children. Longitudinal examination found variability of the strength of the p-factor during the early development, and indicated dynamic maturation, and some support for p-differentiation, as feasible mechanisms of psychopathology development. It also found evidence of genetic influences on p-factor that varied over time. Additionally, the impact of adoptive maternal psychopathology was found to significantly contribute to greater cross-domain difficulties (i.e., higher p-factor scores) in children, as well as specific internalizing and externalizing problems. This present study was limited in that it used a restricted set of items and one informant to model child p-factor. Future research would benefit from using alternative statistical models to complement the current findings, particularly in gathering information of specific symptom pathways and causal mechanisms; and to investigate the role of gene environment interactions and correlations.

References

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 & Profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles.* Burlington, VT.: University of Vermont, Research Center for Children, Youth, & Families.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders.* Arlington, USA: American Psychiatric Publishing.

Andrews, G., & Peters, L. (1998). The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology*. Germany: Springer. http://doi.org/10.1007/s001270050026

B Campbell, S., Spieker, S., Burchinal, M., & Poe, M. (2006). Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*, 791–800. http://doi.org/10.1111/j.1469-7610.2006.01636.x

Bagner, D. M., Pettit, J. W., Lewinsohn, P. M., Seeley, J. R., & Jaccard, J. (2013). Disentangling the temporal relationship between parental depressive symptoms and early child behavior problems: a transactional framework. *Journal of Clinical Child and Adolescent Psychology : The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *42*(1), 78–90. http://doi.org/10.1080/15374416.2012.715368

Beck, C.T. (1999). Maternal depression and child behaviour problems: a meta-analysis. *Journal of Advanced Nursing*, *29*(3), 623–629. http://doi.org/10.1046/j.1365-2648.1999.00943.x

Beck, A. T., & Steer, R. A. (1990). *Beck Anxiety Inventory manual.* San Antonio, Texas: The Psychological Corporation.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, (56), 893–897.

Beck, A. T., & Steer, R. A. (1987). *Manual for the Beck Depression Inventory.* San Antonio, Texas: The Psychological Corporation.

Beck, A. T., Steer, R. A., & Garbin, G. M. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, *8*, 77–100. http://doi.org/10.1016/0272-7358(88)90050-5

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood. *Twin Research and Human Genetics*, *10*, 423–433.

Black, M. M., Jukes, M. C. H., & Willoughby, M. T. (2017). Behavioural and emotional problems in preschool children. *The Lancet Psychiatry*, *4*(2), 89–90. http://doi.org/10.1016/S2215-0366(17)30005-6

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three Concerns With Applying a Bifactor Model as a Structure of Psychopathology. *Clinical Psychological Science*, *5*(1), 184–186. http://doi.org/10.1177/2167702616657069

Brikell, I., Larsson, H., Lu, Y., Pettersson, E., Chen, Q., Kuja-Halkola, R., … Martin, J. (2018). The contribution of common genetic risk variants for ADHD to a general factor of childhood psychopathology. *Molecular Psychiatry*.

http://doi.org/10.1038/s41380-018-0109-2

Brodzinsky, D. M. (2011). Children's Understanding of Adoption: Developmental and Clinical Implications. *Professional Psychology: Research and Practice*, *42*(2), 200–207. http://doi.org/10.1037/a0022415

Brown, T. A. (2014). *Confirmatory Factor Analysis for Applied Research, Second Edition.* New York, NY:Guilford Publications.

Cadore, R., Yates, W., Ed, T., Woodworth, G., & Stewart, M. (1995). Adoption study demonstrating two genetic pathways to drug abuse. *Archives of General Psychiatry*, *52*(1), 42–52.

Capaldi, D. M., & Stoolmiller, M. (1999). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: III. Prediction to young-adult adjustment. *Development and Psychopathology*, *11*(1), 59–84.

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *32*(7), 1063–1080.

Carver, C. S., Johnson, S. L., & Timpano, K. R. (2017). Toward a Functional View of the p Factor in Psychopathology. *Clinical Psychological Science*, *5*(5), 880–889. http://doi.org/10.1177/2167702617710037

Caspi, A., Houts, R. M., Belsky, D. W., & Goldman-mellor, S. J. (2014). The p Factor : One General Psychopathology Factor in the. *Clinical Psychological Science*, *2*(2), 119–137. http://doi.org/10.1177/2167702613497473.

Caspi, A., Moffitt, T., A, T., Freedman, D., Amell, J., Harrington, H., … LA., S. P. (1996). The Life History Calendar: A Research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, *6*(2), 101–114.

Conger, R. D., Ge, X., Elder, G. H., Lorenz, F. O., & Simons, R. L. (1994). Economic Stress, Coercive Family Process, and Developmental Problems of Adolescents. *Child Development*, *65*, 541–561.

Creamer, M., Foran, J., & Bell, R. (1995). The Beck Anxiety Inventory in a non-clinical sample. *Behaviour Research and Therapy*, *33*(4), 477–485. http://doi.org/10.1016/0005-7967(94)00082-U

E. Moffitt, T. (2005). Genetic and Environmental Influences on Antisocial Behaviors: Evidence from Behavioral-Genetic Research. *Advances in Genetics*, *55*(5), 41–104. http://doi.org/10.1016/S0065-2660(05)55003-X

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Eley, T. C., Bolton, D., O'Connor, T. G., Perrin, S., Smith, P., & Plomin, R. (2003). A twin study of anxiety-related behaviours in pre-school children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(7), 945–960. http://doi.org/10.1111/1469-7610.00179

Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal Depression and Child Psychopathology: A Meta-Analytic Review. *Clinical Child and Family Psychology Review*, *14*(1), 1–27. http://doi.org/10.1007/s10567-010-0080-1

Hasin, D., & Kilcoyne, B. (2012). Comorbidity of psychiatric and substance use disorders in the United States: current issues and findings from the NESARC. *Current Opinion in Psychiatry*, *25*(3), 165–171. http://doi.org/10.1097/YCO.0b013e3283523dcc

Hettema, J. M., Neale, M. C., & Kendler, K. S. (2001). A Review and Meta-Analysis of the Genetic Epidemiology of Anxiety Disorders. *American Journal of Psychiatry*, *158*(10), 1568–1578. http://doi.org/10.1176/appi.ajp.158.10.1568

Hu, L.-T., & Bentler, P. . (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

Hudziak, J. J., van Beijsterveldt, C. E., Bartels, M., Rietveld, M. J., Rettew, D. C., Derks, E. M., & Boomsma, D. I. (2003). Individual differences in agression: Genetic analysis by age, gender and informant in 3-, 7- and 10-year old Dutch twins. *Behaviour Genetics*, *33*(5), 575–589. http://doi.org/10.1023/A:1025782918793

Kendler, K. ., Gardner, C. O., & Prescott, C. . (2002). Toward a comprehensive developmental model for major depression in women. *American Journal of Psychiatry*, *159*, 1133–1145.

Kendler, K. S., Prescott, C. a, Myers, J., & Neale, M. C. (2003). The Structure of Genetic and Environmental Risk Factors for Common Psychiatric and Substance Use Disorders in Men and Women, *Archives of General Psychiatry 60* (9)*,* 929-37.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602. http://doi.org/10.1001/archpsyc.62.6.593

Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, *13*(2), 93–121. http://doi.org/10.1002/mpr.168

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, *2*, 111–133.

http://doi.org/10.1146/annurev.clinpsy.2.022305.095213

Lacourse, E., Boivin, M., Brendgen, M., Petitclerc, A., Girard, A., Vitaro, F., … Tremblay, R. E. (2014). A longitudinal twin study of physical aggression during early childhood: evidence for a developmentally dynamic genome. *Psychological Medicine*, *44*(12), 2617–2627. http://doi.org/10.1017/S0033291713003218

Laflamme, D., Pomerleau, A., & Malcuit, G. (2002). A Comparison of Fathers' and Mothers' Involvement in Childcare and Stimulation Behaviors During Free-Play with Their Infants at 9 and 15 Months. *Sex Roles*, *47*(11), 507–518. http://doi.org/10.1023/A:1022069720776

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is There a General Factor of Prevalent Psychopathology during Adulthood? *Journal of Abnormal Psychology*, *121*(4), 971–977. http://doi.org/10.1037/a0028355

Lahey, B. B., Van Hulle, C. A., Singh, A. L., Waldman, I. D., & Rathouz, P. J. (2011). Higher-order genetic and environmental structure of prevalenct forms of child and adolescent psychopathology. *Archives of General Psychiatry*, *68*(2), 181–189.

Leibenluft, E. (2017). Irritability in children: what we know and what we need to learn. *World Psychiatry*, *16*(1), 100–101. http://doi.org/10.1002/wps.20397

Leve, L. D., & Neiderhiser, J. M. (2013). The Early Growth and Development Study: a prospective adoption study from birth through middle childhood. *Twin Research and …*, *16*(1), 412–423. http://doi.org/10.1017/thg.2012.126

Leve, L. D., Neiderhiser, J. M., Ge, X., Scaramella, L. V., Conger, R. D., Reid, J. B., … Reiss, D. (2007). The Early Growth and Development Study: A Prospective

Adoption Design. *Twin Research in Human Genetics*, *10*(1), 84–95. http://doi.org/10.1016/j.molcel.2007.05.041.A

Leve, L. D., DeGarmo, D. S., Bridgett, D. J., Neiderhiser, J. M., Shaw, D. S., Harold, G. T., … Reiss, D. (2013). Using an Adoption Design to Separate Genetic, Prenatal, and Temperament Influences on Toddler Executive Function. Developmental Psychology, 49(6), 1045–1057. http://doi.org/10.1037/a0029390

Lewis, G. J., & Plomin, R. (2015). Heritable influences on behavioural problems from early childhood to mid-adolescence: evidence for genetic stability and innovation. *Psychological Medicine*, *45*(10), 2171–2179. http://doi.org/10.1017/S0033291715000173

Low, S. M., & Stocker, C. (2005). Family Functioning and Children's Adjustment: Associations Among Parents' Depressed Mood, Marital Hostility, Parent-Child Hostility, and Children's Adjustment. Journal of Family Psychology, 19(3), 394-403. http://dx.doi.org/10.1037/0893-3200.19.3.394

Marceau, K., Hajal, N., Leve,L.D., Reiss, D., Shaw, D. S., Ganiban, J.M., Mayes, L. C., Neiderhiser, J. M. (2013). Measurement and associations of pregnancy risk factors with genetic influences, postnatal environmental influences, and toddler behavior. *International Journal of Behavioral Development. 37* (4), 366 - 375. https://doi.org/10.1177/0165025413489378

Marceau, K., De Araujo-Greecher, M., Miller, E. S., Massey, S. H., Mayes, L. C., Ganiban, J. M., … Neiderhiser, J. M. (2016). The perinatal risk index: Early risks experienced by domestic adoptees in the United States. *PLoS ONE*, *11*(3), 1–11. http://doi.org/10.1371/journal.pone.0150486

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., … Salum, G. A. (2017). A general psychopathology factor (P Factor) in

children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology*, *126*(1), 137–148. http://doi.org/10.1037/abn0000205

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology*, *22*(3), 491–495. http://doi.org/10.1017/S0954579410000222

Mcelroy, E., Belsky, J., Carragher, N., Fearon, P., & Patalay, P. (2017). Developmental stability of general and specific factors of psychopathology from early childhood to adolescence: Dynamic mutualism or p-differentiation? *Journal of Child Psychology and Psychiatry and Allied Disciplines*. http://doi.org/10.1111/jcpp.12849

McElroy, E., Fearon, R. M. P., Belsky, J., Fonagy, P., & Patalay, P. (in press). Networks of depression and anxiety symptoms across development. *Journal of the American Academy of Child and Adolescent Psychiatry.*

Melby, J. N., Conger, R. D., Ge, X., & Warner, T. (1995). The use of structural equation modeling in assessing the quality of marital observations. *Journal of Family Psychology*, *9*, 280–293.

Merikangas, K. R., He, J.-P., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001-2004 NHANES. *Pediatrics*, *125*(1), 75–81. http://doi.org/10.1542/peds.2008-2598

Moffitt, T. E. (2005). The New Look of Behavioral Genetics in Developmental Psychopathology: Gene-Environment Interplay in Antisocial Behaviors. *Psychological Bulletin*. *131* (4), 533-54. http://doi.org/10.1037/0033-2909.131.4.533

Mott, S. L., Schiller, C. E., Richards, J. G., O'Hara, M. W., & Stuart, S. (2011).

Depression and anxiety among postpartum and adoptive mothers. *Archives of Women's Mental Health*, *14*(4), 335. http://doi.org/10.1007/s00737-011-0227-1

Murray, A. L., Eisner, M., & Ribeaud, D. (2016). The Development of the General Factor of Psychopathology "p Factor" Through Childhood and Adolescence. *Journal of Abnormal Child Psychology*, *44*(8), 1573–1586. http://doi.org/10.1007/s10802-016-0132-1

Murray, A. L., McKenzie, K., Kuenssberg, R., & O'Donnell, M. (2014). Are we under-estimating the association between autism symptoms?: The importance of considering simultaneous selection when using samples of individuals who meet diagnostic criteria for an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(11), 2921–2930. http://doi.org/10.1007/s10803-014-2154-2

Muthén, L. K., & Muthén, B. O. (2017). *Mplus User's Guide. Eighth Edition.* (8th ed.). Los Angeles, California: Muthén & Muthén.

Neumann, A., Pappa, I., Lahey, B. B., Verhulst, F. C., Medina-Gomez, C., Jaddoe, V. W., … Tiemeier, H. (2016). Single Nucleotide Polymorphism Heritability of a General Psychopathology Factor in Children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *55*(12), 1038–1045.e4. http://doi.org/10.1016/j.jaac.2016.09.498

Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology*. US: American Psychological Association. http://doi.org/10.1037/0021-843X.107.2.305

Nivard, M. G., Dolan, C. V., Kendler, K. S., Kan, K. J., Willemsen, G., Van Beijsterveldt, C. E. M., … Boomsma, D. I. (2015). Stability in symptoms of anxiety and depression as a function of genotype and environment: A

longitudinal twin study from ages 3 to 63 years. *Psychological Medicine*, *45*(5), 1039–1049. http://doi.org/10.1017/S003329171400213X

O'Connor, T. G., Neiderhiser, J. M., Reiss, D., Hetherington, E. M., & Plomin, R. (1998). Genetic contribution to continuity, change, and co-occurrence of antisocial and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry*, *39*(3), 323–336.

Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *British Journal of Psychiatry*, *207*(1), 15–22. http://doi.org/10.1192/bjp.bp.114.149591

Patterson, G., & Stoolmiller, M. (1991). *Replications of a dual failure model for boys' depressed mood*. *Journal of consulting and clinical psychology* (Vol. 59). http://doi.org/10.1037/0022-006X.59.4.491

Pechtel, P., & Pizzagalli, D. (2012). Effects of Early Life Stress on Cognitive and Affective Function. *Psychopharmacology*, *214*(1), 55–70. http://doi.org/10.1007/s00213-010-2009-2.Effects

Pemberton, C. K., Neiderhiser, J. M., Leve, L. D., Natsuaki, M. N., Shaw, D. S., Reiss, D., & Ge, X. (2010). Influence of Parental Depressive Symptoms on Adopted Toddler Behaviors: An Emerging Developmental Cascade of Genetic and Environmental Effects. *Development and Psychopathology*, *22*(4), 803–818. http://doi.org/10.1017/S0954579410000477

Pettersson, E., Larsson, H., & Lichtenstein, P. (2015). Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. *Molecular Psychiatry*, *21*(5), 717-721.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological*

Bulletin. US: American Psychological Association. http://doi.org/10.1037/0033-2909.84.2.309

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, *88*(5), 879–903. http://doi.org/10.1037/0021-9010.88.5.879

Polderman, T. J. C., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*(7), 702–709. http://doi.org/10.1038/ng.3285

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). NIH Public Access. *Journal of Personality Assessment*, *92*(6), 544–559. http://doi.org/10.1080/00223891.2010.496477.

Reise, S. P., & Waller, N. G. (2009). Item Response Theory and Clinical Measurement. *Annual Review of Clinical Psychology*, *5*(1), 27–48. http://doi.org/10.1146/annurev.clinpsy.032408.153553

Rhee, S. H., & Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin*. Rhee, Soo Hyun: U Colorado, Inst for Behavioral Genetics, Campus Box 447, Boulder, CO, US, 80309, soo.rhee@colorado.edu: American Psychological Association. http://doi.org/10.1037/0033-2909.128.3.490

Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Archives of General Psychiatry*, *38*(4), 381–389.

Robins, L. N., Helzer, J. E., Ratcliff, K. S., & Seyfried, W. (1982). Validity of the Diagnostic Interview Schedule, .2. Dsm-Iii Diagnoses. *Psychol Med*, *12*(4), 855–870.

Rutter, M. (2006). *Genes and Behavior: Nature-Nurture Explained*. Oxford, UK: Blackwell Publising Ltd.

Rutter, M., Bishop, D., Pine, D., Scott, S., Stevenson, J. S., Taylor, E., & Thapar, A. (2008). *Rutter's Child and Adolescent Psychiatry*. http://doi.org/10.1002/9781444300895

Salvatore, J. E., & Dick, D. M. (2016). Genetic influences on conduct disorder. *Neuroscience and Behavioral Reviews*, *In Press*(804). http://doi.org/10.1016/j.neubiorev.2016.06.034

Scaini, S., Belotti, R., & Ogliari, A. (2014). Genetic and environmental contributions to social anxiety across different ages: A meta-analytic approach to twin data. *Journal of Anxiety Disorders*, *28*(7), 650–656. http://doi.org/https://doi.org/10.1016/j.janxdis.2014.07.002

Scarr, S., & McCartney, K. (1983). How people make their own environments: a theory of genotype greater than environment effects. *Child Development*, *54*(2), 424–435.

Schumann, L., Boivin, M., Paquin, S., Lacourse, E., Brendgen, M., Vitaro, F., … Booij, L. (2017). Persistence and innovation effects in genetic and environmental factors in negative emotionality during infancy: A twin study. *PLoS ONE*, *12*(4), e0176601. http://doi.org/10.1371/journal.pone.0176601

Senecky, Y., Agassi, H., Inbar, D., Horesh, N., Diamond, G., Bergman, Y. S., & Apter, A. (2009). Post-adoption depression among adoptive mothers. *Journal of Affective Disorders*, *115*(1), 62–68. http://doi.org/10.1016/j.jad.2008.09.002

Shimada-Sugimoto, M., Otowa, T., & Hettema, J. M. (2015). Genetics of anxiety disorders: Genetic epidemiological and molecular studies in humans. *Psychiatry and Clinical Neurosciences*, *69*(7), 388–401. http://doi.org/10.1111/pcn.12291

Smith, D. W., & M., B. D. (2002). Coping with birthparent loss in adopted children. *Journal of Child Psychology and Psychiatry*, *43*(2), 213–223. http://doi.org/10.1111/1469-7610.00014

Smoller, J. W., Andreassen, O. A., Kendler, K. S., Edenberg, H. J., Faraone, S. V, & Glatt, S. J. (2018). Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry*, 1. http://doi.org/10.1038/s41380-017-0010-4

Snyder, H. R., Young, J. F., & Hankin, B. L. (2017). Strong Homotypic Continuity in Common Psychopathology-, Internalizing-, and Externalizing-Specific Factors Over Time in Adolescents. *Clinical Psychological Science*, *5*(1), 98–110. http://doi.org/10.1177/2167702616651076

Spielberger, C. D., Gorsuch, R. L., & R.E., L. (1970). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, California: Consulting Psychologists Press.

Stover, C. S., Connell, C. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Scaramella, L. V., … Reiss, D. (2012). Fathering and mothering in the family system: Linking marital hostility and aggression in adopted toddlers. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(4), 401–409. http://doi.org/10.1111/j.1469-7610.2011.02510.x

Sullivan, P. F., Neale, M. C., & Kendler, K. S. (2000). Genetic Epidemiology of Major Depression: Review and Meta-Analysis. *American Journal of Psychiatry*, *157*(10), 1552–1562. http://doi.org/10.1176/appi.ajp.157.10.1552

Tackett, J. L., Lahey, B. B., van Hulle, C., Waldman, I., Krueger, R. F., & Rathouz,

P. J. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology*, *122*(4), 1142–1153. http://doi.org/10.1037/a0034151

Tieger, T. (1980). On the biological basis of sex differences in aggression. *Child Development*, *51*(4), 943–963.

van Bork, R., Epskamp, S., Rhemtulla, M., Borsboom, D., & van der Maas, H. L. J. (2017). What is the p-factor of psychopathology? Some risks of general factor modeling. *Theory and Psychology*, *27*(6), 759–773. http://doi.org/10.1177/0959354317737185

van den Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology*. US: American Psychological Association. http://doi.org/10.1037/0021-843X.105.3.349

van der Maas, H. L. J., Dolan, C. V, Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological Review*, *113*(4), 842–861. http://doi.org/10.1037/0033-295X.113.4.842

Van Houtem, C. M. H. H., Laine, M. L., Boomsma, D. I., Ligthart, L., van Wijk, A. J., & De Jongh, A. (2013). A review and meta-analysis of the heritability of specific phobia subtypes and corresponding fears. *Journal of Anxiety Disorders*, *27*(4), 379–388. http://doi.org/https://doi.org/10.1016/j.janxdis.2013.04.007

Waldman, I. D., Poore, H. E., van Hulle, C., Rathouz, P. J., & Lahey, B. B. (2016). External validity of a hierarchical dimensional model of child and adolescent psychopathology: Tests using confirmatory factor analyses and multivariate behavior genetic analyses. *Journal of Abnormal Psychology*, *125*(8), 1053–1066. http://doi.org/10.1037/abn0000183

Willner, C. J., Gatzke-Kopp, L. M., & Bray, B. C. (2016). The dynamics of internalizing and externalizing comorbidity across the early school years. *Development and Psychopathology*, *28*(4pt1), 1033–1052. http://doi.org/10.1017/S0954579416000687

Part 3: Critical appraisal

Introduction

In some ways the experience of working on a secondary data project can feel quite detached from the traditional client-facing clinical role. However, it was my personal interest in the field of developmental psychopathology and the excitement to be able to ask “big questions” about the aetiology of psychopathology and the influences of early environment on development that led me to choosing this current project. Yet when describing my decision to others, some of the reactions and opinions were that this type of project was more suited and relevant to “academics” than practicing psychologists. More specifically their reservations centred around the value of the skills obtained from working on secondary data projects and the relevance of genetics to our practice.

Consequently, this critical appraisal will address some of my reflections about why psychologists should become involved in secondary data projects; how issues relating to the conceptualisation of psychopathology and genetics are relevant to our current practice; and finally why it’s important that clinical psychologists remain interested and involved in these areas.

Why should psychologists work on secondary data projects?

By working on a large scale secondary data project , the experience provides an invaluable reminder that, whilst large datasets may provide us with more confidence in answers to questions that smaller scale research may not, it is not without its own set of limitations (Schofield, 2017). The questions you are able to ask are limited to the data that has been collected already and the quality of the questionnaires used. There are the questions of how to manage missing data and at each stage pragmatic decisions are made that will ultimately impact the results and limit the interpretations that can be made. In the current climate, national healthcare services are influenced largely by the findings from routinely collected outcome measures (Gyani, Shafran, Layard, & Clark, 2011), and there is increasing

recognition of the potential of “big data” collected by mental health records to help in the development of the delivery of mental health services (McIntosh et al., 2016; Raghupathi & Raghupathi, 2014). Thus, experience in working with large datasets is becoming increasingly relevant to our field.

“Big data” projects, also known as data science, refers to the use of advanced statistical methods to extract knowledge from large complex datasets (McIntosh et al., 2016). Whilst the statistical analyses employed in working with big data require a more advanced application of statistical skills and knowledge (e.g. structural equation modelling), even without those skills our training as psychologists places us at an advantage in understanding and translating the findings from these types of projects. Particularly for big data projects based on routine data collection, clinical psychologists are well positioned to understand and provide a sensible interpretation of the results from their own clinical experiences of working in services. For example, knowledge of the on-the-ground human processes that occur in data collection, e.g. confirmatory biases, team values around data collection, etc. It is predicted that such large scale projects may impact service provision and decisions on healthcare, the consequences of which will be great given the limited resources that face the UK mental health service (British Medical Association, 2017). Therefore it is important that clinical psychologists are involved at the very least in conversations related to the findings of these projects, to ensure that findings and interpretations are applied critically and interpreted within in the context of what also isn’t known (Schofield, 2017).

Furthermore, working with big data could highlight one possible way of providing more information about this area where there is a lack of funding for expensive RCTs (Fischer et al., 2013) . This is particularly relevant for the field of mental health research, that comparatively receives less funding than its’s physical health counterparts (Kingdon, 2006). Currently, in the structure of service provision,

both academically and clinically, there is little priority given to preventing mental health problems by intervening during the very early stages in life (British Medical Association, 2017). This is despite the knowledge that half of individuals that develop enduring mental health problems will do so before the age of 14 (Kessler et al., 2005) and that emotional and behaviour problems can be identified from very early ages. Whilst preventing mental health problems for infants and pre-schoolers appears to be higher in the research agendas in countries outside of the UK (such as what was found in the review from Part 1), more attention and research is still needed to bridge existing gaps in our knowledge.

How do we conceptualise mental health problems?

Given the potential implications and likely future use of big data to inform service provision and policy, it raises some important points for consideration about what should be prioritised in research. Whilst treatment efficacy and outcome measurement remain important, the focus on how we conceptualise mental health problems and the development of psychopathology needs also to remain firm on research agendas and in the interest of clinicians. Indeed the focus on the aetiology of mental health problems is as important now as it ever was, with increasing rates of mental health difficulties reported in children (Bor, Dean, Najman, & Hayatbakhsh, 2014).

As outlined in the empirical paper, the current psychiatric nosology systems offer problematic conceptualisations of psychopathology that present challenges to, and have potentially limited, our current understanding of the development of mental health problems. Poor definitions of mental health problems subsequently mean flawed interventions, but also poor measures of outcomes that feed into service provision (Wolpert & Rutter, 2018). If there is no good measure for mental health problems then it raises questions as to what is already known about mental health, and leaves the effectiveness of our own interventions uncertain.

Furthermore, a recent longitudinal prospective cohort study in New Zealand, the Dundein Study, has shown that it is the norm for most people to experience acute episodes of mental health problems during their lifetime (Schaefer et al., 2017). They followed individuals from birth to midlife and found only 17% prevalence rate for those that had never met criteria a mental health disorder (Schaefer et al., 2017). Arguably the current provision of psychotherapy in services is based on the conceptualised model of the treatment of acute mental health problems. Whilst their findings need replicating, it raises some interesting points for consideration: if the experience of mental health problems is potentially the normality, it highlights a need to better identify and understand the risk mechanisms behind those that go on to develop enduring mental health problems. Having a better system for identifying and conceptualising those most at risk for enduring mental health difficulties is paramount, particularly if we hope to improve our understanding of the aetiological mechanisms involved in this profile, which will ultimately lead to improved interventions.

Why should clinical psychologists care about genetics?

Moreover, in our pursuit to further understand the aetiology of psychopathology, our attention towards biopsychosocial processes should not be neglectful to the role of biological influences. Certainly in my own experience, it is more common for psychologists to consider the psychological and social explanations and understandings of the development and maintenance of different disorders, with less attention being paid to the role of underlying biological factors, for example the implications in treatment of genetic vulnerabilities.

Certainly the field of behavioural genetics has a controversial and appalling history which may contribute to some of the lack of enthusiasm, or caution in the integration of findings from this field of research. Additionally, the narrow pursuit of finding a gene “for” psychiatric disorders in the deterministic sense, which is often

much of the focus from the media, does not help in the appearance of clinically relevant findings for the psychologist. Yet, the evidence provided from this field undeniably shows that genetics have an important role to play in our understanding of the development of psychopathology. Indeed, it has been shown that genes operate to influence behaviour in many ways both directly and indirectly through interaction with the environment (Rutter, 2006). Through genetically sensitive study designs, the interaction of the individual with the environment can be examined and further insight can be shed on mitigating and aggravating risk factors for developing psychopathology. This has direct and important consequences for advancing our own scientific understanding within our field, which can be integrated with our current theories and models and also improve the clinical interventions in our practice.

There is a need to understand the casual developmental cascades of those that experience enduring mental health problems and through improving our understanding of genetic risks it could provide a way to identify children most at risk for future difficulties and assist our understanding of likely prognosis for that child which could help prevent future difficulties. Indeed, should research improve identification of specific genes for psychopathology, it may be that genetic counselling would be beneficial in identifying those that need interventions the most, i.e. are the most genetically vulnerable (Thapar & Rutter, 2009). Additionally, better understanding of the dynamic interplay between genetics and environmental factors would further our understanding of what types of intervention would be helpful to administer (i.e. best environmental risk to target such as parental mental health problems), and when the timing of intervention may be the most effective (i.e. if there are particularly vulnerable times in development).

What do families and practitioners think about genetics?

Genetic research has illustrated that bi-directional influences between child (their genetics) and their environment (namely their parents) are the rule rather than the exception (Thapar & Rutter, 2009). Despite this, from my experience, often the primacy of the environment is valued by clinicians. For some parents this can leave them feeling blamed or responsible for the emotional and behavioural problems of their child. Even for some families that were engaging with services, they were desperate for a preferential diagnoses of ADHD or autism as an explanation for their child's behaviour, which they saw as more of a "biological" problem which would be less blaming towards themselves. It's important that psychologists help to address this perceived stigma, and ensure the role of genetics is not neglected in shared formulation with families about emotional and behavioural problems. This is particularly important for child services as it may help in improving engagement, before the development of more serious problems.

Following on from this, it is particularly important for practitioners to consider the role of genetics in work with adoptive families. The knowledge of the birth parents psychiatric history may not always be known by adoptive parents, and even when their history is known, it is important to clarify their understanding of what genetics means to them. For example, it is important that parents understand that whilst most behavioural traits show some modest heritability (Rutter, 2006), genes indicate probabilistic risks and do not mean that the child will deterministically develop the same problems (of their birth parents, for example).

It would also be helpful for practitioners to think with adoptive parents about the how different genetic profiles between themselves and the adopted child may interact in terms of their combined temperaments and how that may influence the effectiveness of parenting strategies. For example, a child may be more behaviourally inhibited and want to stay in close proximity with the parent when in

new situations; for more outgoing parents the child's wariness may cause frustration, yet thinking about these differences in terms of differences in biological or temperamental dispositions may foster a more empathetic understanding of the difficulties. Whilst this is the same in many ways for parents of biological children, for some families of adopted children little is known about the child's earlier years and it may raise more questions about whether it is the child's nature or the impact of trauma that is causing a child to behave in such ways. Therefore, discussions about the role of nature (genetics) and the responses to the environment may be even more important.

Conclusion

In conclusion, working with secondary data has challenged many of my own reservations regarding the relevance of this type of project for a clinical psychologist. It has made me appreciate the value of clinical psychologists engaging in in interdisciplinary research, particularly for behavioural genetics, and the importance of ensuring that findings from academic studies don't become isolated from clinical practice. Particularly as psychologists are in a position of being able to shape public discourse about the development of mental health problems through our clinical work in teams and with clients.

References

Bor, W., Dean, A. J., Najman, J., & Hayatbakhsh, R. (2014). Are child and adolescent mental health problems increasing in the 21st century? A systematic review. *Australian & New Zealand Journal of Psychiatry*, *48*(7), 606–616. http://doi.org/10.1177/0004867414533834

British Medical Association. (2017). *Breaking down barriers – the challenge of improving mental health outcomes*.

Fischer, A. J., Threlfall, A., Meah, S., Cookson, R., Rutter, H., & Kelly, M. P. (2013). The appraisal of public health interventions: An overview. *Journal of Public Health (United Kingdom)*, *35*(4), 488–494. http://doi.org/10.1093/pubmed/fdt076

Gyani, A., Shafran, R., Layard, R., & Clark, D. (2011). Enhancing Recovery Rates in IAPT Services: lessons from analysis of the Year One data.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602. http://doi.org/10.1001/archpsyc.62.6.593

Kingdon, D. (2006). Health research funding. *BMJ*, *332*(7556), 1510. Retrieved from http://www.bmj.com/content/332/7556/1510.1.abstract

McIntosh, A. M., Stewart, R., John, A., Smith, D. J., Davis, K., Sudlow, C., … Porteous, D. J. (2016). Data science for mental health: a UK perspective on a global challenge. *The Lancet Psychiatry*, *3*(10), 993–998. http://doi.org/10.1016/S2215-0366(16)30089-X

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, *2*(1), 3.

http://doi.org/10.1186/2047-2501-2-3

Rutter, M. (2006). *Genes and Behavior: Nature-Nurture Explained*. Oxford, UK: Blackwell Publising Ltd.

Rutter, M., Bishop, D., Pine, D., Scott, S., Stevenson, J. S., Taylor, E., & Thapar, A. (2008). *Rutter's Child and Adolescent Psychiatry*. http://doi.org/10.1002/9781444300895

Schaefer, J. D., Caspi, A., Belsky, D. W., Harrington, H., Houts, R., Horwood, L. J., … Moffitt, T. E. (2017). Enduring mental health: Prevalence and prediction. *Journal of Abnormal Psychology*, *126*(2), 212–224. http://doi.org/10.1037/abn0000232

Schofield, P. (2017). Big data in mental health research – do the *n* s justify the means? Using large data-sets of electronic health records for mental health research. *BJPsych Bulletin*, *41*(3), 129–132. http://doi.org/10.1192/pb.bp.116.055053

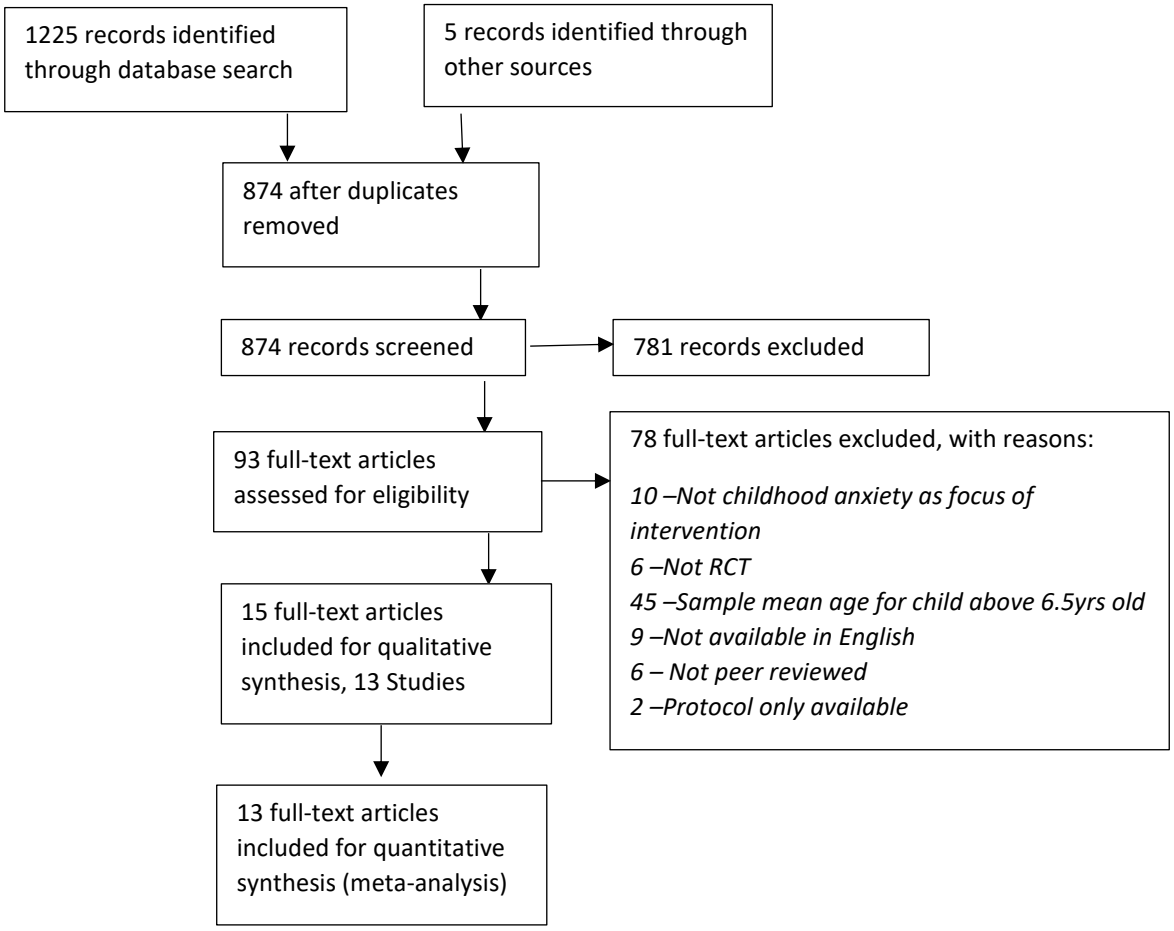
Thapar, A., & Rutter, M. (2009, February 10). Genetics. *Rutter's Child and Adolescent Psychiatry*. http://doi.org/doi:10.1002/9781444300895.ch23

Wolpert, M., & Rutter, H. (2018). Using flawed , uncertain , proximate and sparse (FUPS) data in the context of complexity : learning from the case of child mental health, 1–11. http://doi.org/10.1186/s12916-018-1079-6

Appendices

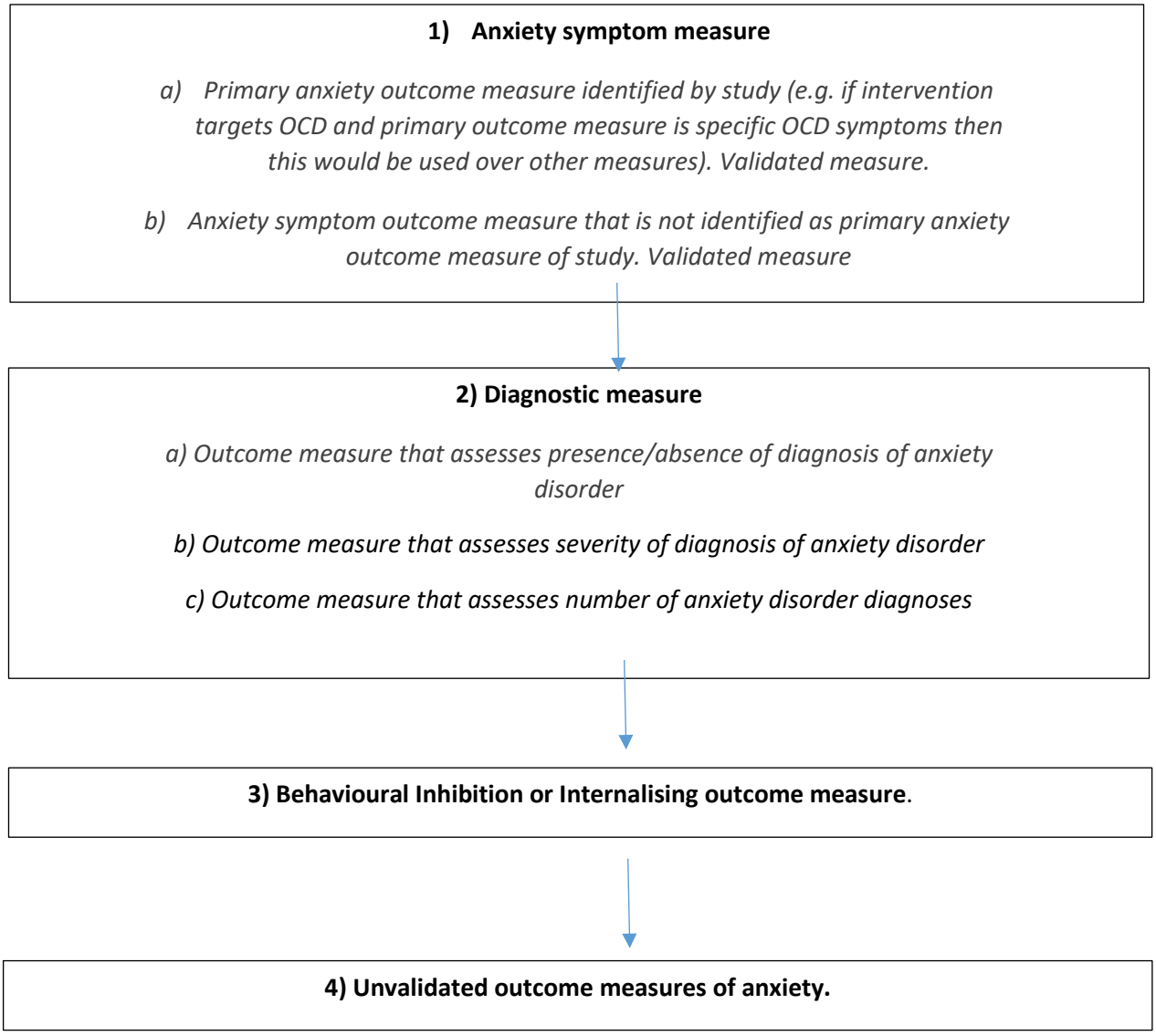
Appendix A

Figure 1. PRISMA flowchart



Appendix B

- Figure 2. Figure illustrating decision hierarchy for selection of measures used in meta-analysis



Appendix C

Table 1 .Table of included studies with outcome measures and key findings for anxiety measures.

Author (year)	Anxiety Measures	Key findings for anxiety measures	Follow-up included
Schneider, Blatter-Meunier, Herren, Adornetto, In-Albon, Lavallee (2001)	• Kinder-DIPS (Schnieder, Unnewhr & Margraf, 2009) • **Separation Anxiety Inventory for Children (SAI; Scalbert, In-Albon & Schnieder, 2006)** • Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1978)	Results indicated significantly greater reduction of separation anxiety post-treatment for IG, compared to CG.	Yes, for all groups. Treatment gains were maintained at the 4-week follow-up, 76.19% of children in IG no longer met diagnostic criteria for SAD (compared to 13.64% in the CG).
Sanacruz, Mendez& Sanchez-Meca (2008)	• Dark Fear Interview (DFI; Méndez, 1996) • Children's Fear Survey Schedule-Revised (CFSS-R; Pelechano, 1984) • Dark Fear Scale (DFS; Méndez&Santacruz,1996) • **Dark Behavior Recording-Modified (DBR-M; Mikulas & Coffman, 1989)**	Both play therapies achieved a significant reduction in darkness phobia compared to CG.	Yes, for all groups. Treatment gains were maintained at 12-month follow up.
Rapee, Kennedy, Ingram, Edwards & Sweeneey (2005; 2010)	• Laboratory assessment for Behaviour Inhibition (Kagan et al., 1989, 1994) • Short Temperament Scale for Children (STSC), which is an abbreviated version of the Childhood Temperament Questionnaire (Australian version; Sanson, Smart, Prior, Oberklaid, & Pedlow, 1994; Thomas & Chess, 1977). • Temperament Assessment Battery for Children—Revised (TABC–R; Presley & Martin, 1994).	Children whose parents were allocated to the education condition showed a significantly greater decrease in anxiety diagnoses at 12 months relative to those whose parents received no intervention. However, there were no significant effects demonstrated on measures of inhibition/withdrawal.	Yes, for all groups. For anxiety diagnosis and measures there were no significant main effects of group, however there were group and time effects reported at 36-month follow up. There were lower levels of children's self-reported symptoms of anxiety in the IG compared to the CG, although this

	• Anxiety Disorders Interview Schedule for Children and Parents IV—Parent Version (Silverman & Albano, 1996). • **Spence Children's Anxiety Scale (Spence, Rapee, McDonald & Ingram, 2001)**		was not statistically significant.
Pahl & Barrett (2010)	• **Preschool Anxiety Scale, Parent Report (Spence et al, 2001)** • Behavioural Inhibition Questionnaire (Bishop et al., 2003)	Parent report data revealed no significant differences between the IG and CG on anxiety and behavioural inhibition at post-intervention.	Yes, for IG only. Improvements were found on anxiety and BI measures for IG at 12-month follow-up.
Morgan, Rapee, Tamir, Goharpey, Salim, McLellan & Bayer (2015;2017)	• Short Temperament Scale for Children (STSC; Prior, Sanson & Obkerlaid,1989) • Strengths and Difficulties Questionnaire – Parent version (SDQ-P; Goodman, 1997) • **Revised Preschool Anxiety Scale (PAS-R; Edwards, Rapee, Kennedy & Space, 2010).** • Online Assessment of Preschool Anxiety (OAPA)	The IG showed significantly greater improvement over time in child anxiety symptoms and lower rates of anxiety disorders compared to the CG.	Yes, for all groups. There was a significantly greater improvement on anxiety symptoms at 24-week follow-up for IG compared to CG.
Menzies & Clarke (1993)	• **Behaviour Rating Scale (BRS; Willis, 1983)** • Water phobia survey schedule (Willis, 1983) • Overall reaction to phobic situation	Both IVVE and IVE conditions led to significant treatment gains. However, the VE condition did not lead to substantial improvement from pre- to post-treatment.	Yes, for IGs only. At 12-week follow-up only IVVE group showed slight improvement on measures. IVE showed poorer maintenance of treatment gains. On average follow-up scores across IGs were slightly higher than those at post-treatment, though this finding was not significant.

Lewin et al. (2014)	• Anxiety Disorders Interview Schedule eParent Version (ADIS; Silverman & Albano, 1996) • **Children Yale-Brown Obsessive Compulsive Scale (CYBOCS; Scahill et al., 1997)** • Pediatric Anxiety Rating Scale (PARS; RUPP, 2002) • Clinical Global Impression -Severity and -Improvement (CGI-S/-I; National Institute of Mental Health, 1985) • National Institute of Mental Health Global OCD Scale (NIMH-GOCS; Goodman & Price, 1992)	There was a large main effect of group for anxiety measures for the IG. 65% of the IG were classified as treatment responders as compared to 7% in the CG.	Yes, for IG only. Treatment gains for IG were maintained at three month follow-up.
Kennedy, Rapee & Edwards (2009)	• **Preschool Anxiety Scale (PAS-R; Spence et al, 2001)** • Disorders Interview Schedule for Children and Parents IV-Parent Version (ADIS-IV-P; Silverman & Albano, 1996) • Laboratory assessment for Behaviour Inhibition (Kagan et al., 1989, 1994) • Short Temperament Scale for Children (STSC; Prior, Sanson & Obkerlaid,1989) • Behavioural Inhibition Questionnaire (BIQ; Bishop, Spence, McDonald & Ingram, 2003)	Compared to the CG, the IG showed a significantly greater reduction in anxiety disorders. For measures of BI, IG showed largest reductions.	No, six-month follow-up was only post-intervention measure.
Hirshfeld-Becker, Masek & Henin (2010)	• **Schedule for Affective Disorders and Schizophrenia, Epidemiologic Version (K-SADS-E; Orvaschel, 1994)** • CGI- Anxiety 7-Point rating • Adapted laboratory assessment for Behaviour Inhibition (Rosenbaum et al., 2000) • Child Behaviour Checklist (CBCL; Achenbach, 1991)	IG showed a significantly greater reduction in anxiety disorders and significantly better CGI improvement on social phobia/avoidant disorder, separation anxiety disorder and specific phobia (but not generalised anxiety disorder) compared to CG.	Yes, for IG only. Treatment gains were maintained at one year follow-up.

Donovan & March (2014)	• Anxiety Disorders Interview Schedule for DSM-IV: Parent version (ADIS-P; Silverman & Albano, 1996) • **Preschool Anxiety Scale (PAS; Spence, Rapee, McDonald, & Ingram, 2001).** • Child Behaviour Checklist (CBCL; Achenbach, 1991)	IG at post-treatment showed a significantly greater reduction in anxiety symptoms, clinical severity and internalising behaviour compared to CG. There were no significant group differences for the percentages of children who lost their primary anxiety diagnosis or who lost all anxiety diagnoses.	Yes, for IG only. Treatment gains were maintained for overall functioning and further improved with respect to anxiety symptoms, clinical severity and internalising behaviour at six-month follow-up.
Anticich, Barrett, Silverman, Lacherez & Gillies (2013)	• **Preschool Anxiety Scale (PAS; Spence et al., 2001)** • Behavioural Inhibition Questionnaire (BIQ; Bishop, Spence, McDonald & Ingram, 2003)	Results were comparable for IGs and CG, however the 'Fun FRIENDS' IG achieved greater reductions in BI.	Yes for all groups. Both IGs improved significantly more than CG, with 'Fun FRIENDS' IG improving significantly more than 'You Can Do it' IG at 12-month follow-up.
Waters, Ford, Wharton & Cobham (2009)	• **Anxiety Disorders Interview Schedule for the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV): Child/Parent Versions (ADIS-C-IV-C/P; Silverman & Albano, 1996)** • Spence Children's Anxiety Scale – Parent version (SCAS-P; Nauta et al., 2004) • Child Behaviour Checklist (CBCL; Achenbach & Eldelbrock,1983; Achenbach & Rescorla, 2000)	Both IGs were superior to the CG, with 54.8% of children in the 'Parent & Child' and 55.3% of children in the 'Parent Only' IG no longer meeting criteria for their primary diagnosis at post-treatment.	Yes, for IG only. Treatment gains were maintained in both IGs at six-month and 12-month follow-up assessments.
Cartwright-Hatton, McNally, Field,	• **Screen for Child Anxiety Related Disorders-Parent version (SCARED;**	Children from IG were 7.35 time	Yes, for all groups. Both groups demonstrated a

Rust, Laskey & Dixon et al. (2011)	**Birmaher, Khetarpal, Cully, Brent & McKenzie, 1997)** • Multidimensional Anxiety Scale for Children (MASC; March, Parker, Sullivan, Stallings & Connor, 1997) • Anxiety Disorders Interview Schedule for Children and Parents-IV-Parent-version (ADIS-PV; Silverman & Albano, 1996)	s more likely to be free of primary diagnosis than CG. Decrease in total SACRED and internalising scores was greater in IG compared to CG.	reduction in symptoms in self-report measures of anxiety at 12-month follow-up, however the reduction in scores by IG was not significantly stronger than CG. For anxiety diagnosis IG 8.5 times more likely to be free of primary diagnosis than CG.

Note: Items marked in bold indicate measure chosen for meta-analysis. IG=Intervention Group; CG= Control Group.

Appendix D

Figure 1. Bi-factor model containing preschool item loadings onto the internalising and externalising specific factors and the general psychopathology bi-factor (p-factor).

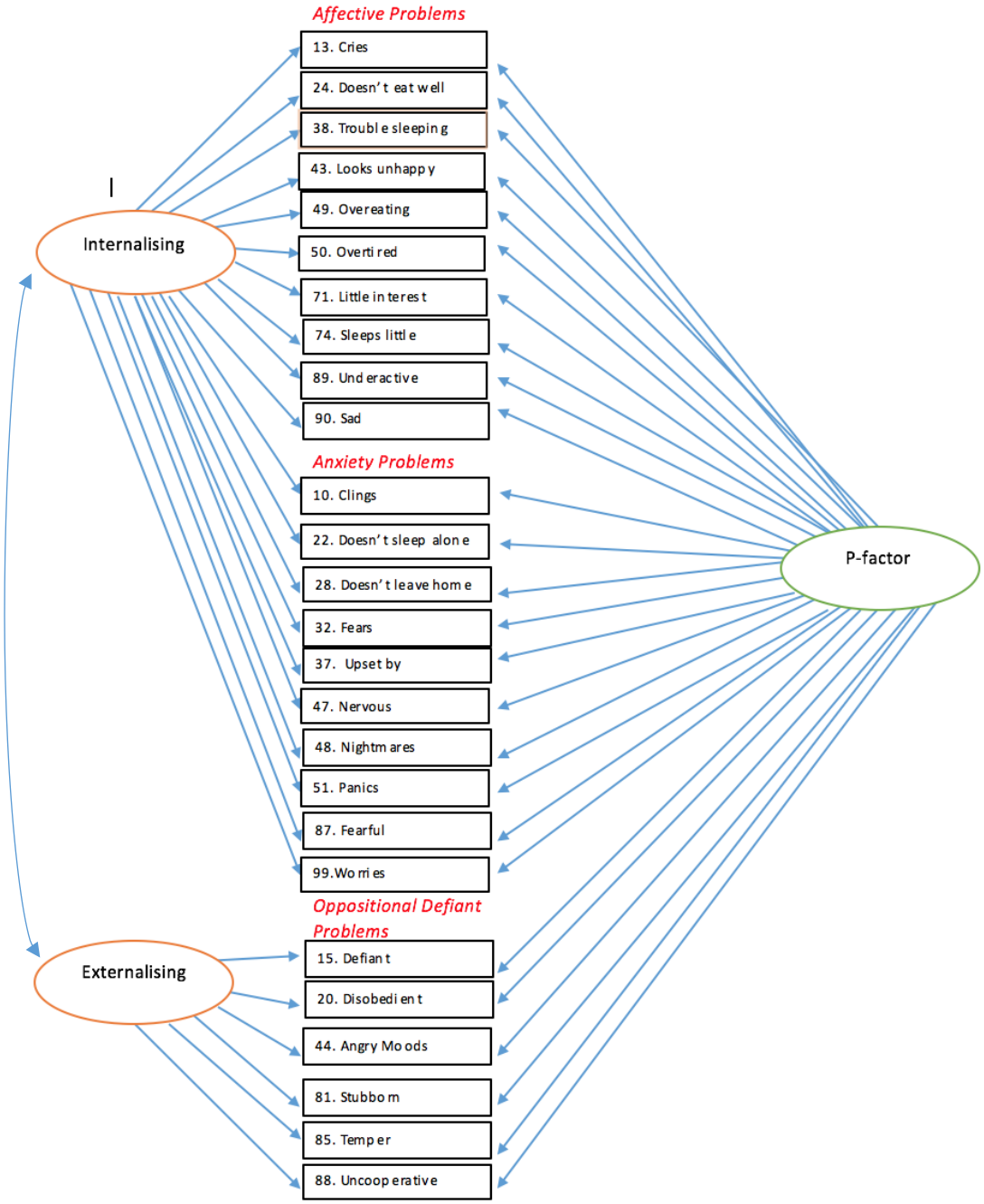
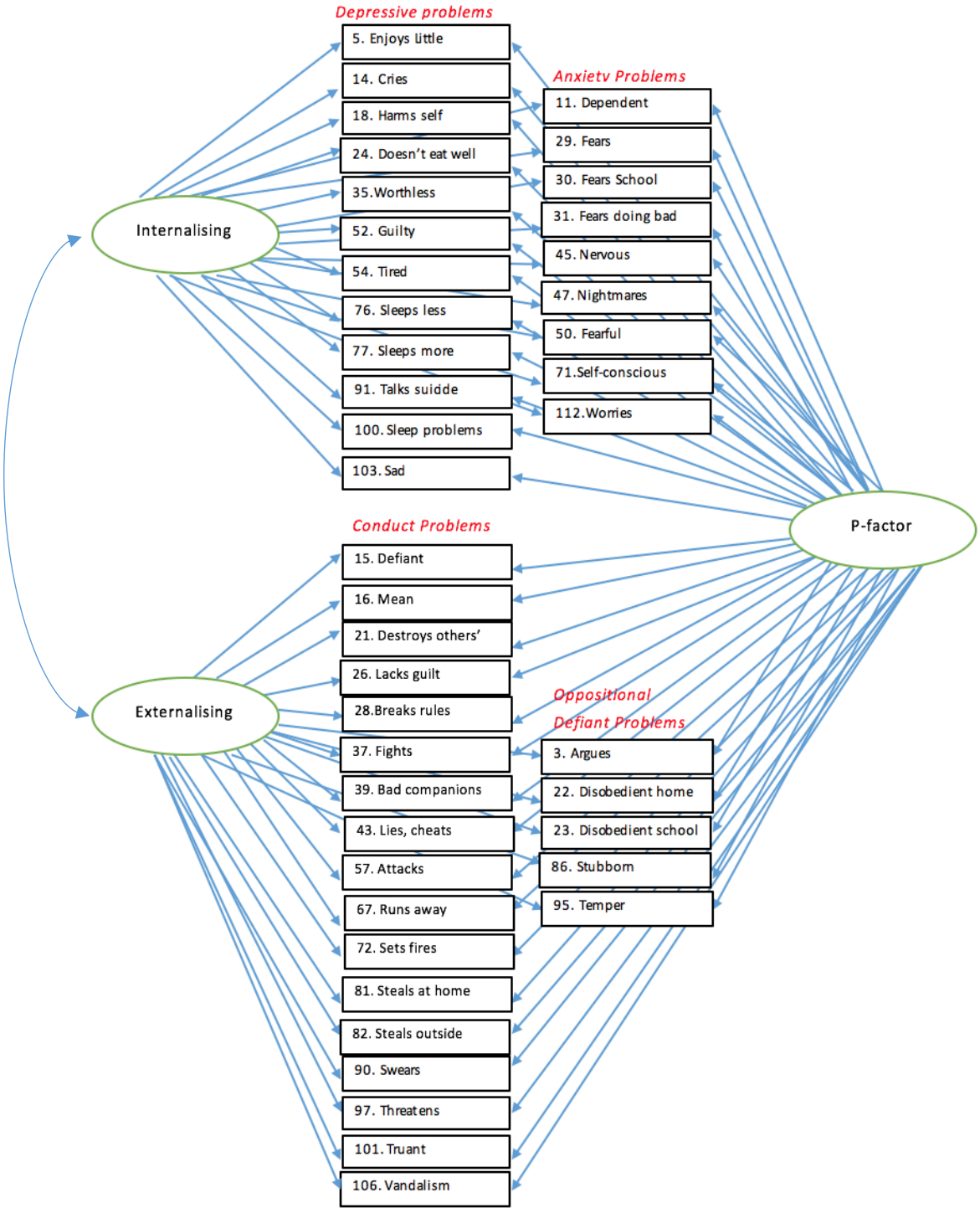
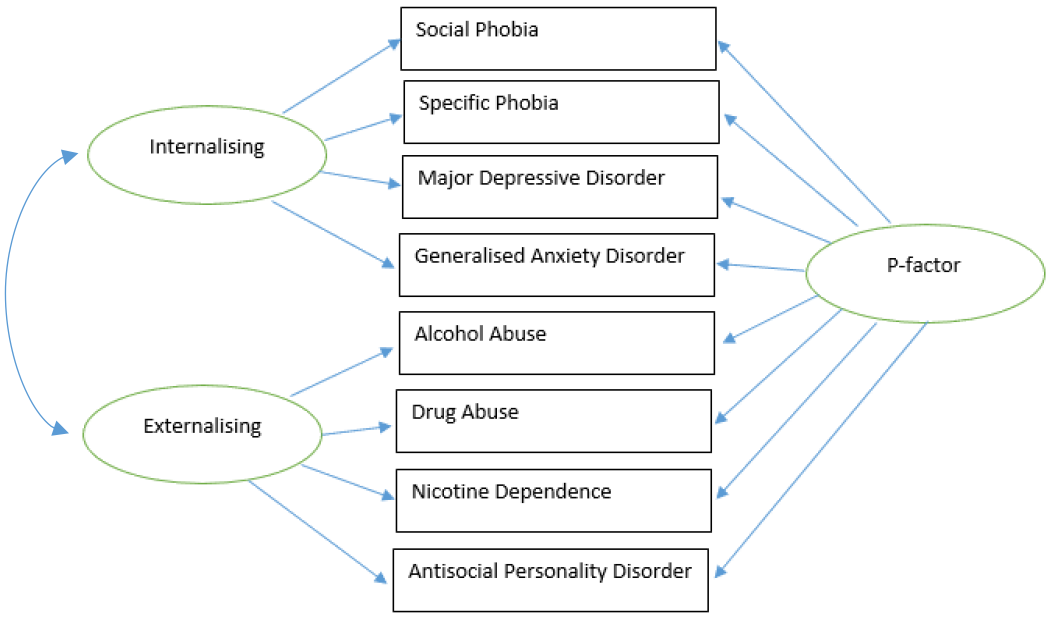


Figure 2. Bi-factor model containing school age item loadings onto the internalising and externalising specific factors and the general psychopathology bi-factor (p-factor).



Appendix E

Figure 3. Figure showing bi-factor model containing DSM loadings onto the internalising and externalising specific factors and the general psychopathology bi-factor (p-factor)



Appendix F

Table 5. Table summarising Explained Common Variance for correlated (CFA) and p-factor (CBM) model over time.

	CFA		**CBM**		
	Internalising	*Externalising*	*Internalising*	*Externalising*	*P-factor*
18 months	0.39	0.61	0.02	0.48	0.51
27 months	0.37	0.63	0.22	0.59	0.20
4.5 years	0.39	0.61	0.01	0.56	0.43
7 years	0.31	0.69	0.00	0.05	0.95

Appendix G

Table 6. Standardised CBCL item factor loadings for correlated (CFA) and p-factor (CBM) model at 18 months.

	CFA		CBM		
Factor loadings	Internalising	Externalising	Internalising	Externalising	P
13. Cries	0.59		0.08 *(ns)*		0.56
24. Doesn't eat well	0.27		0.15 *(ns)*		0.19
38. Trouble Sleeping	0.52		0.80		0.19
43. Looks unhappy	0.87		-0.03 *(ns)*		0.84
49. Overeating	0.22		0.11 *(ns)*		0.19 *(ns)*
50. Overtired	0.56		0.41		0.39
71. Little interest	-		-		-
74. Sleeps little	0.52		0.62		0.23
89. Underactive	-		-		-
90. Sad	-		-		-
10. Clings	0.34		0.27		0.23
22. Doesn't sleep alone	0.34		0.69		-0.02 *(ns)*
28. Doesn't leave home	-		-		-
32. Fears	0.39		0.12 *(ns)*		0.34
37. Upset by separation	0.32		0.35		0.16
47. Nervous	-		-		-
48. Nightmares	0.39		0.36		0.24
51. Panics	-		-		-
87. Fearful	-		-		-
99. Worries	-		-		-
15. Defiant		0.74		0.54	0.56
20. Disobedient		0.82		0.63	0.62
44. Angry Moods		0.72		0.03 *(ns)*	0.76
81. Stubborn		0.75		0.17*(ns)*	0.76
85. Temper		0.6		0.21*(ns)*	0.57
88. Uncooperative		0.84		0.49	0.69

INT= internalising; EXT=externalising. All factor loadings except the ones marked *ns* are significant at least at the 0.05.

Table 7. Standardised CBCL item factor loadings for correlated (CFA) and p-factor (CBM) model at 27 months.

Factor loadings	CFA		CBM		
	INT	*EXT*	*INT*	*EXT*	*P*
13. Cries	0.63		0.47		0.44
24. Doesn't eat well	0.22		0.22		0.07 *(ns)*
38. Trouble Sleeping	0.33		0.59		-0.31 *(ns)*
43. Looks unhappy	0.72		0.32 *(ns)*		0.75
49. Overeating	-		-		-
50. Overtired	0.57		0.65		0.06 *(ns)*
71. Little interest	-		-		-
74. Sleeps little	0.38		0.61		-0.28
89. Underactive	-		-		-
90. Sad	-		-		-
10. Clings	0.55		0.44		0.33
22. Doesn't sleep alone	0.22		0.48		-0.38
28. Doesn't leave home	0.44		0.43		0.14 *(ns)*
32. Fears	0.44		0.25		0.4
37. Upset by separation	0.43		0.35		0.25
47. Nervous	-		-		-
48. Nightmares	0.29		0.31		0.06 *(ns)*
51. Panics	-		-		-
87. Fearful	0.7		0.21 *(ns)*		0.81
99. Worries	0.61		0.17 *(ns)*		0.73
15. Defiant		0.81		0.78	0.24 *(ns)*
20. Disobedient		0.85		0.86	0.18 *(ns)*
44. Angry Moods		0.80		0.65	0.5
81. Stubborn		0.78		0.71	0.32
85. Temper		0.7		0.59	0.37
88. Uncooperative		0.76		0.75	0.19 *(ns)*

INT= internalising; EXT=externalising. All factor loadings except the ones marked *ns* are significant at least at the 0.05.

Table 8. Standardised CBCL item factor loadings for correlated (CFA) and p-factor (CBM) models at 4.5 years.

Factor loadings	CFA		CBM		
	INT	*EXT*	*INT*	*EXT*	*P*
13. Cries	0.59		0.08 *(ns)*		0.58
24. Doesn't eat well	0.3		-0.01 *(ns)*		0.29
38. Trouble Sleeping	0.41		-0.33		0.44
43. Looks unhappy	0.76		0 *(ns)*		0.78
49. Overeating	0.41		-0.17 *(ns)*		0.43
50. Overtired	0.61		-0.26		0.65
71. Little interest	-		-		-
74. Sleeps little	0.49		-0.23 *(ns)*		0.51
89. Underactive	0.59		0.06 *(ns)*		0.58
90. Sad	0.71		-0.2 *(ns)*		0.73
10. Clings	0.44		0.45		0.39
22. Doesn't sleep alone	0.35		-0.09 *(ns)*		0.35
28. Doesn't leave home	0.57		0.29		0.54
32. Fears	0.35		0.55		0.29
37. Upset by separation	0.50		0.36		0.46
47. Nervous	0.72		0.14 *(ns)*		0.71
48. Nightmares	0.47		-0.13 *(ns)*		0.49
51. Panics	0.74		0.33		0.69
87. Fearful	0.76		0.72		0.65
99. Worries	0.59		0.34		0.55
15. Defiant		0.74		0.67	0.43
20. Disobedient		0.81		0.85	0.39
44. Angry Moods		0.74		0.35	0.61
81. Stubborn		0.74		0.33	0.62
85. Temper		0.78		0.46	0.6
88. Uncooperative		0.85		0.67	0.56

INT= internalising; EXT=externalising. All factor loadings except the ones marked *ns* are significant at least at the 0.05.

Table 9. Standardised CBCL item factor loadings for correlated (CFA) and p-factor (CBM) models at 7 years.

Factor loadings	**CFA**		**CBM**		
	INT	*EXT*	*INT*	*EXT*	*P*
5. Enjoys little	0.47		0.06 *(ns)*		0.45
14. Cries	0.47		0.15 *(ns)*		0.4
18. Harms self	-		-		-
24. Doesn't eat well	0.4		0.08		0.359
35. Worthless	0.76		0.38		0.592
52. Guilty	-		-		-
76. Sleeps less	0.48		0.43		0.287
77. Sleeps more	-		-		-
91. Talks suicide	-		-		-
100. Sleep problems	0.57		0.38		0.4
103. Sad	0.78		0.37		0.62
11. Dependent	0.61		0.56		0.35
29. Fears	0.37		0.55		0.1 *(ns)*
30. Fears School	0.49		0.6		0.2
31. Fears doing bad	0.57		0.4		0.39
45. Nervous	0.71		0.58		0.43
47. Nightmares	0.5		0.24		0.39
50. Fearful	0.72		0.87		0.3
71. Self-conscious	0.61		0.3		0.48
112. Worries	0.69		0.61		0.42
3. Argues		0.69		0.09 *(ns)*	0.69
22. Disobedient home		0.86		0.34	0.8
23. Disobedient school		0.64		0.57	0.53
86. Stubborn		0.82		0.01 *(ns)*	0.83
95. Temper		0.78		-0.24 *(ns)*	0.85
15. Defiant		-		-	-
16. Mean		0.77		0.18 *(ns)*	0.74
21. Destroys others'		0.68		0.13 *(ns)*	0.67
26. Lacks guilt		0.56		0.35	0.5
28. Breaks rules		0.83		0.5	0.74

37. Fights	-	-	-
39. Bad companions	0.44	0.22 *(ns)*	0.4
43. Lies, chats	0.73	0.38	0.67
57. Attacks	0.79	-0.04*(ns)*	0.8
67. Runs away	-	-	-
72. Sets fires	-	-	-
81. Steals at home	-	-	-
82. Steals outside	-	-	-
90. Swears	-	-	-
97. Threatens	0.8	-0.08 *(ns)*	0.81
101. Truant	-	-	-
106. Vandalism	-	-	-

INT= internalising; EXT=externalising. All factor loadings except the ones marked *ns* are significant at least at the 0.05

Appendix H

Table 10. Table showing standardised DSM diagnosis factor loadings for correlated (CFA) and p-factor (CBM) models for birth mother.

Factor loadings	CFA		CBM		
	INT	*EXT*	*INT*	*EXT*	*P*
Social Phobia	0.64		0.28 *(ns)*		0.59
Specific Phobia	0.37		0.76 *(ns)*		0.23 *(ns)*
Major Depressive Disorder	0.70		0.27*(ns)*		0.64
Generalised Anxiety Disorder	0.92		0.08 *(ns)*		0.92
Alcohol Abuse		0.66		0.52	0.35
Drug Abuse		0.80		0.89	0.36
Nicotine Dependence		0.62		0.29	0.54
ASPD		0.62		0.45	0.39

INT= internalising; EXT=externalising; ASPD= Antisocial Personality Disorder. All factor loadings except the ones marked *ns* are significant at least by the 0.05

Appendix I

Table 13. Table showing a summary of examination of evocative rGE using Pearson's correlation to examine the genetic influence of the birth mother's bi-factor model on environment measures.

	N	*BM p-factor*	BM externalising	BM internalising
		18 months		
AM psych	473	.00	-.07	.02
AF psych	450	.01	-.02	.04
Economic distress	485	-.06	-.04	.00
Marital Hostilty	468	.07	-.01	.06
		27 months		
AM psych	468	.03	-.06	.06
AF psych	482	.00	-.03	.05
Economic distress	498	-.03	.01	-.04
Marital Hostilty	497	.03	-.04	.05
		4.5 years		
AM psych	505	.03	-.07	.06
AF psych	486	.00	-.03	.05
Economic distress	500	-.02	.02	-.04
Marital Hostilty	501	.01	-.06	.02
		7 years		
AM psych	510	.02	**-.09***	.04
AF psych	494	.01	-.04	.03
Economic distress	506	-.02	.01	-.06
Marital Hostilty	507	.01	-.07	.00

*AM psych= adoptive mother psychopathology; AF adoptive father psychopathology; BM=birth mother; *p<.05.*

Appendix J

Table 14. Summary of the parametric and bootstrapped hierarchical regression models of covariate and genetic predictors for the internalising specific factor (from the child bi-factor model) over time.

					Bootstrap			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
18 months (N= 470)									
Openness[1]	.01	.01	.04	.27	0	0	.27	0	.01
Perinatal risk[1]	0	.01	.03	.86	0	0	.87	0	0
Gender[1]	.02	.03	.04	.68	0	.01	.66	-.01	.01
BMint[2]	-.03	.1	-.01	.39	-.02	.02	.38	-.06	.02
BMext[2]	.04	.05	.05	.38	.01	.01	.39	-.01	.03
BMp[2]	0	.04	0	.94	0	.01	.94	-.02	.02
Δ R^2	.01								
F	.77			.51					
27 months (N=447)									
Openness[1]	0	.02	-.01	.83	0	.02	.83	-.04	.03
Perinatal risk[1]	.02	.01	.06	.24	.02	.01	.24	-.01	.04
Gender[1]	.04	.03	.06	.23	.04	.03	.24	-.03	.11
BMint[2]	-.09	.13	-.03	.51	-.09	.13	.51	-.35	.19
BMext[2]	.01	.06	.01	.89	.01	.06	.91	-.11	.14
BMp[2]	.03	.05	.04	.48	.03	.05	.48	-.05	.13
Δ R^2	0								
F	.34			.8					
4.5 years (N=379)									
Openness[1]	0	0	-.03	.52	0	0	.5	-.01	0
Perinatal risk[1]	0	0	0	.94	0	0	.94	0	0
Gender[1]	0	.01	.03	.52	0	.01	.51	-.01	.02
BMint[2]	.01	.02	.02	.75	.01	.02	.76	-.03	.05
BMext[2]	-.02	.01	-.08	.14	-.02	.01	.13	-.04	0
BMp[2]	.01	.01	.08	.14	.01	.01	.12	0	.03
Δ R^2	.01								
F	1.53			.21					
7 years (N=374)									
Openness[1]	0	0	-.09	.07	0	0	.06	-.01	0

Perinatal risk[1]	0	0	-.01	.88	0	0	.89	0	0
Gender[1]	.01	0	.06	.25	.01	0	.26	0	.01
BMint[2]	0	.02	.01	.86	0	.02	.85	-.03	.04
BMext[2]	0	.01	-.01	.92	0	.01	.92	-.02	.02
BMp[2]	0	.01	-.04	.48	0	.01	.48	-.02	.01
Δ R^2	0								
F	.21			.89					

SE*B*= standard error *B;CI=confidence interval*;. AM= adoptive mother; AF=adoptive father; [1]=entered in block one of hierarchical regression;[2]=entered in block two. Significant results are highlighted in bold.

Table 15. Summary of the parametric and bootstrapped hierarchical regression models of covariate and genetic predictors for the externalising specific factor (from the child bi-factor model) over time.

					Bootstrap			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
18 months (N= 470)									
Openness[1]	.01	.01	.04	.41	.01	.01	.35	-.01	.04
Perinatal risk[1]	.01	.01	.03	.47	.01	.01	.49	-.01	.03
Gender[1]	.02	.03	.04	.37	.02	.03	.33	-.03	.07
BMint[2]	-.02	.10	-.01	.81	-.02	.1	.81	-.22	.17
BMext[2]	.04	.05	.04	.38	.04	.05	.38	-.05	.15
BMp[2]	.00	.04	0	.94	0	.04	.92	-.08	.07
Δ R^2	0								
F	.37			.78					
27 months (N=447)									
Openness[1]	-.04	.03	-.06	.2	-.04	.03	.22	-.10	.03
Perinatal risk[1]	0	.02	0	.97	.00	.02	.96	-.04	.04
Gender[1]	.01	.06	.01	.9	.01	.06	.9	-.1	.11
BMint[2]	.02	.22	.01	.91	.02	.23	.92	-.40	.45
BMext[2]	.15	.10	.08	.15	.15	.10	.13	-.04	.34
BMp[2]	.09	.08	.06	.26	.09	.08	.23	-.06	.25
Δ R^2	.01								
F	1.62			.18					
4.5 years (N=379)									
Openness[1]	.04	.03	.07	.15	.04	.02	.12	-.01	.08
Perinatal risk[1]	.01	.02	.03	.54	.01	.02	.54	-.03	.05
Gender[1]	-.07	.05	-.08	.12	-.07	.05	.12	-.17	.02
BMint[2]	.11	.18	.03	.55	.11	.17	.52	-.26	.46
BMext[2]	.09	.09	.06	.28	.09	.09	.29	-.08	.27
BMp[2]	.00	.07	.00	.95	0	.06	.95	-.13	.13
Δ R^2	0								
F	.49			.69					
7 years (N=374)									
Openness[1]	0	0	.06	.22	0	0	.2	0	.01

Perinatal risk[1]	0	0	.04	.40	0	0	.39	0	.01
Gender[1]	**-.02**	**.01**	**-.18**	**<.001**	**-.02**	**.01**	**<.001**	-.03	-.01
BMint[2]	.01	.02	.04	.53	.01	.02	.53	-.03	.05
BMext[2]	0	.01	-.01	.89	0	.01	.9	-.02	.02
BMp[2]	.01	.01	.04	.46	.01	.01	.44	-.01	.02
Δ R^2	0								
F	.48			.69					

AM= adoptive mother; AF=adoptive father*;* SE*B*= standard error *B*; [1]=entered in block one of hierarchical regression;[2]=entered in block two. Significant results are highlighted in bold.

Appendix K

Table 16. Summary of the parametric and bootstrapped hierarchical regression models of environmental predictors for the internalising specific factor over time.

					Bootstrap			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
				18 months (N= 461)					
Openness[1]	0	0	.06	.19	0	0	.21	0	.01
OCT risk[1]	0	0	.01	.8	0	0	.79	0	0
Gender[1]	0	.01	-.03	.58	0	.01	.56	-.01	.01
AM psychopathology[2]	0	0	.06	.21	0	0	.19	.00	.01
AFpsychopathology[2]	0	0	-.02	.72	0	0	.71	-.01	.01
Economic distress[2]	0	0	-.01	.77	0	0	.79	-.01	0
Marital Hostility[2]	**0**	**0**	**.11**	**.03**	**0**	**0**	**.02**	0	0
Δ R^2	.02								
F	2.02			.09					
				27 months (N=459)					
Openness[1]	.01	.02	.02	.60	.01	.02	.62	-.03	.05
OCT risk[1]	.01	.01	.04	.43	.01	.01	.40	-.01	.03
Gender[1]	.04	.03	.05	.24	.04	.03	.24	-.02	.10
AM psychopathology[2]	**.09**	**.02**	**.20**	**<.001**	**.09**	**.02**	**<.001**	.05	.13
AFpsychopathology[2]	-.02	.02	-.05	.34	-.02	.02	.29	-.05	.01
Economic distress[2]	.02	.01	.05	.27	.02	.01	.23	-.01	.04
Marital Hostility[2]	**0**	**0**	**.1**	**.03**	**.00**	**.00**	**.02**	.00	.01
Δ R^2	**.07**								
F	**8.07**			**<.001**					
				4.5 years (N=394)					
Openness[1]	0	0	-.08	.14	0	0	.13	-.01	.00
OCT risk[1]	0	0	0	.94	0	0	.95	.00	.00
Gender[1]	0	.01	0	.95	0	.01	.95	-.01	.01
AM psychopathology[2]	0	0	-.01	.88	0	0	.90	-.01	.01
AFpsychopathology[2]	0	0	-.02	.68	0	0	.69	-.01	.01
Economic distress[2]	0	0	-.06	.28	0	0	.34	-.01	.00
Marital Hostility[2]	0	0	-.05	.36	0	0	.36	.00	.00
Δ R^2	.01	.							
F	86			.49					
				7 years (N=392)					

Openness[1]	**-.01**	**0**	**-.12**	**.02**	**-.01**	**0**	**.02**	-.01	.00
Perinatal risk[1]	0	0	-.03	.6	0	0	.59	.00	.00
Gender[1]	0	0	.06	.25	0	0	.26	.00	.01
AM psychopathology[2]	**.01**	**0**	**.16**	**.003**	**.01**	**0**	**.01**	.00	.02
AFpsychopathology[2]	0	0	.04	.42	0	0	.43	.00	.01
Economic distress[2]	0	0	0	.98	0	0	.98	.00	.00
Marital Hostility[2]	0	0	.04	.41	0	0	.44	.00	.00
Δ R^2	**.04**								
F	**3.85**			**.004**					

AM= adoptive mother; AF=adoptive father; SE*B*= standard error *B*; [1]=entered in block one of hierarchical regression;[2]=entered in block two. Significant results are highlighted in bold.

Table 17. Summary of the parametric and bootstrapped hierarchical regression models of environmental predictors for the externalising specific factor over time.

					Bootstrap			*95% CI*	
	B	SE *B*	β	*p*	*B*	*SE B*	*p*	Lower	Upper
18 months (N= 461)									
Openness[1]	0	.01	.01	.75	0	.01	.74	-.02	.03
Perinatal risk[1]	0	.01	.01	.77	0	.01	.77	-.02	.02
Gender[1]	.02	.03	.03	.52	.02	.03	.52	-.03	.06
AM psychopathology[2]	**.04**	**.02**	**.12**	**.02**	**.04**	**.02**	**.02**	.01	.08
AFpsychopathology[2]	-.02	.02	-.08	.12	-.02	.02	.15	-.05	.01
Economic distress[2]	0	.01	-.02	.72	0	.01	.7	-.03	.02
Marital Hostility[2]	0	0	-.03	.54	0	0	.53	.00	.00
Δ R^2	.02								
F	2.02			.09					
27 months (N=459)									
Openness[1]	-.01	.03	-.01	.78	-.01	.03	.78	-.07	.05
Perinatal risk[1]	0	.02	0	.95	0	.02	.95	-.04	.04
Gender[1]	.03	.06	.03	.55	.03	.05	.57	-.08	.14
AM psychopathology[2]	**.17**	**.04**	**.22**	**<.001**	**.17**	**.04**	**<.001**	.10	.23
AFpsychopathology[2]	-.01	.03	-.01	.81	-.01	.03	.78	-.07	.05
Economic distress[2]	.04	.02	.08	.09	.04	.02	.06	0	.09
Marital Hostility[2]	0	0	.01	.81	0	0	.8	-.01	.01
Δ R^2	**.06**								
F	**7.8**			**<.001**					
4.5 years (N=394)									
Openness[1]	**.05**	**.03**	**.11**	**.04**	**.05**	**.02**	**.04**	.01	.1
Perinatal risk[1]	.02	.02	.05	.3	.02	.02	.3	-.01	.05
Gender[1]	-.07	.05	-.08	.1	-.07	.05	.1	-.16	.01
AM psychopathology[2]	**.08**	**.03**	**.14**	**.01**	**.08**	**.03**	**.01**	.02	.14
AFpsychopathology[2]	-.03	.03	-.04	.40	-.03	.03	.34	-.09	.02
Economic distress[2]	0	.02	0	.96	0	.02	.96	-.04	.04
Marital Hostility[2]	-.01	0	-.09	.09	-.01	0	.09	-.01	0

Δ R^2	.02								
F	2.17			.07					
			7 years (N=392)						
Openness[1]	0	0	.08	.12	0	0	.09	0	.01
Perinatal risk[1]	0	0	.02	.71	0	0	.7	0	0
Gender[1]	**-.02**	**.01**	**-.15**	**.002**	**-.02**	**.01**	**.01**	-.03	-.01
AM psychopathology[2]	.01	0	.09	.09	.01	0	.07	0	.01
AFpsychopathology[2]	0	0	-.05	.33	0	0	.32	-.01	0
Economic distress[2]	0	0	.02	.7	0	0	.7	0	.01
Marital Hostility[2]	0	0	-.07	.19	0	0	.18	0	0
Δ R^2	.01								
F	1.32			.26					

Appendix L

Figure 6. Graph illustrating estimated sample size needed to detect different effect sizes using the Fisher's z test with the reference value constrained to zero. In this study small effect sizes (<.2) may not be detected due to lack of power.