

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**NEUROPSYCHOLOGICAL ASSESSMENT IN MIDDLE CHILDHOOD:
Objective and subjective assessment of executive and social functioning**

A thesis presented in partial fulfilment of the
requirements for the degree of

Doctor of Clinical Psychology

at Massey University, Wellington,
New Zealand.

Kathryn Marie McLennan

2017

ACKNOWLEDGEMENTS

I would like to start by thanking my wonderful supervisor Janet Leathem, whose vision, support, and expertise were invaluable in putting this thesis together. I am very fortunate to have had her in my corner.

I would like to thank Kate and Yanis, whose efforts in data collection were phenomenal, Ross Flett for his support and feedback in the final stages of thesis completion, and my colleagues in the Centre of Public Health Research for their generosity and kindness. I would also like to acknowledge the participating schools, staff and students for being so accommodating and for all their hard work. It was a pleasure working with you.

Finally I would like to thank my friends, my family, and Alex who have never ceased to be supportive even long distance. You have made this process so much easier than it otherwise could have been. Thank you.

ABSTRACT

This thesis presents a research study that aimed to explore measurement issues in child neuropsychological assessment, within a NZ cultural context.

Neuropsychological assessment tools should be developmentally and culturally appropriate, yet most measures used with NZ children have not been evaluated with this population. Further, both subjective and objective assessment tools have been developed, but it is unclear how information gained from these assessment tools relate to each other and inform clinical practice.

Child neuropsychology has undergone many changes in recent years. With an increasing understanding of autism spectrum disorders, new domains of functioning have been introduced in neuropsychological assessment instruments, notably, executive functioning (EF), theory of mind (ToM), and affect recognition (AR). Numerous similarities have been documented between these constructs however, making interpretation of assessment results complex. Though these constructs are thought to develop across middle childhood, this population is vastly understudied.

This thesis aimed to evaluate patterns of performance on the BASC-2, BRIEF and NEPSY-II measures with NZ school-aged children. Normative data for these measures is not available for NZ children, therefore this research aimed to evaluate the suitability of test norms for this demographic. Further, the thesis aimed to explore the relationship between parent and teacher reports of function and evaluate how subjective (broadband and narrowband) and objective measures of EF, ToM and AR are related. Participants were 241 children from schools within the Wellington and Hawke's Bay regions, recruited as part of a larger study.

Ratings on the BASC-2 and BRIEF measures differed substantially from American norms, with parents and teachers tending to report fewer problem behaviours and more adaptive behaviours than their American counterparts. Results indicated a discrepancy between parent and teacher reports, and between the information gained from subjective versus objective measures. The results of this research have important clinical implications not just for the use and interpretation of these measures, but also for the neuropsychological assessment of EF, ToM and AR in NZ children.

TABLE OF CONTENTS

Acknowledgements	ii
Abstract	iii
Table of Contents	v
Table of Tables	x
Table of Figures	xi
Abbreviations	xii
CHAPTER ONE	
PREFACE	1
CHAPTER TWO	
NEUROPSYCHOLOGICAL ASSESSMENT	5
Overview	5
Neuropsychological Assessment	5
Developmental Assessment	5
Cultural Differences in Assessment	6
Assessment Modalities	7
Assessment of Cognition	8
Assessment of Behaviour	9
Summary	11
CHAPTER THREE	
EXECUTIVE FUNCTIONING, THEORY OF MIND AND AFFECT RECOGNITION	12
Overview	12
Executive Function	13
Theory of Mind	15
Affect Recognition (AR)	18
Relationship between the Constructs	19
EF and ToM	19
EF and AR	22
ToM and AR	22
Measurement Issues	23
Cultural appropriateness of EF, ToM and AR measures	23
Individual differences in performance	24
Developmental EF assessment	25
Developmental ToM assessment	26
Overlap in EF and ToM assessment	27
Subjective versus objective assessment	28
Summary	29

CHAPTER FOUR	
THE PROPOSED RESEARCH	30
Gaps in the Current Literature	30
Aims	31
Significance of Research	33
Research Context	34
CHAPTER FIVE	
METHOD	37
Procedure	37
Phase 1	38
Phase 2	39
Participants	39
Measures	43
Behavior Assessment System for children – 2 nd edition (BASC-2)	43
Behaviour Rating Inventory of Executive Function (BRIEF)	46
Neuropsychological Assessment Measures	47
Developmental NEuroPSYchological Assessment (NEPSY-II)	49
Social perception subtests	50
ToM	50
AR	51
EF subtests	51
Inhibition	51
Animal Sorting	51
Stimulus Materials	52
NEPSY-II Scoring	52
Changes to Test Administration	53
Inter-Rater Reliability	54
Exclusion Criteria	54
Testing Environment	54
Compensation	55
Confidentiality	56
Feedback	56
CHAPTER SIX	
PRELIMINARY ANALYSES	57
Overview	57
Data Collation	57
Testing Assumptions of Normality	58
Internal Consistency	59
Examiner Effects	59

CHAPTER SEVEN	
STUDY ONE	61
Overview	61
Part One	61
Overview	61
Results	62
Discussion	65
Part Two	66
Overview	66
Results	66
Evaluation of the BASC-2	66
Evaluation of the BRIEF	69
Evaluation of the NEPSY-II subtests	72
Discussion	73
Part Three	76
Overview	76
Results	76
Gender comparisons	76
Comparisons across decile	79
Discussion	81
Gender comparisons	81
Comparisons across decile	81
CHAPTER EIGHT	
STUDY TWO	83
Overview	83
Results	83
Relationship between parent and teacher responses on the BASC-2	83
Relationship between parent and teacher responses on the BRIEF	86
Discussion	88
Relationship between parent and teacher responses on the BASC-2	88
Relationship between parent and teacher responses on the BRIEF	89
CHAPTER NINE	
STUDY THREE	90
Overview	90
Results	90
Evaluation of BASC-2 ratings with NEPSY-II performance	90
Evaluation of BRIEF ratings with NEPSY-II performance	92

Discussion	93
Evaluation of BASC-2 ratings with NEPSY-II performance	93
Evaluation of BRIEF ratings with NEPSY-II performance	95
CHAPTER TEN	
CONCLUSION	97
Summary of Findings	97
Implications for Clinical Practice	98
Limitations and Future Directions	99
Sample characteristics	99
Decile distribution of the sample	99
Gender and ethnicity distribution of the sample	100
The measures	102
BASC and BRIEF questionnaires	102
NEPSY-II	102
Neuropsychological testing	102
Order effects	102
Assessment period	103
Testing environment	103
Conceptual Framework	104
Avenues for Future Research	105
Overview of Author Reflections	106
Conclusions	107
REFERENCES	108
APPENDIX A	
INFORMATION AND CONSENT FORMS	123
Parent Information Sheet	123
Parent Reply Form	125
Teacher Information Sheet	126
Teacher Consent Form	128
APPENDIX B	
ADAPTIONS TO SELECTED MEASURES	129
Attachment to BASC-2 and BRIEF questionnaires	129
Table B1. List of cultural adaptations made to NEPSY-II ToM task	130
APPENDIX C	
NON-PARAMETRIC STATISTICS	131
Table C1. Non-parametric analyses examining whether NEPSY-II scores differ as a function of examiner	131

Table C2. Non-parametric analyses examining whether BASC-2 and BRIEF scores differ from US norms	131
Table C3. Non-parametric analyses examining whether NEPSY-II scores differ from US norms	132
Table C4. Non-parametric analyses examining whether BASC-2 and BRIEF scores differ across categories of decile	132
Table C5. Non-parametric analyses examining whether BASC-2 and BRIEF scores differ as a function of gender	133

APPENDIX D	
REFLECTIONS ON RESEARCH PROCESS AND ITS APPLICATION TO CLINICAL PRACTICE	134
Abstract	135
Introduction	136
Research Skills	137
Knowledge Gained Through Thesis Research	140
Neuropsychological assessment	140
Cognitive development	141
Working with children	142
Working with schools	143
Working in a team (multi-disciplinary)	144
Other learning	146
Discussion	147
References	148

TABLE OF TABLES

Table 1.	<i>Demographic composition of sample by school</i>	43
Table 2.	<i>Subtests administered during neuropsychological assessment</i>	49
Table 3.	<i>Available n for comparison across measures</i>	58
Table 4.	<i>Mean scores on NEPSY-II subtests by examiner</i>	60
Table 5.	<i>Frequency of PRS and TRS Critical Items</i>	63
Table 6.	<i>One-sample t-tests comparing BASC-2 scores from NZ sample with US test mean of 50 (SD = 10)</i>	68
Table 7.	<i>One-sample t-tests comparing BRIEF scores from NZ sample with US test mean of 50 (SD = 10)</i>	71
Table 8.	<i>One-sample t-tests comparing NEPSY-II scores from NZ sample with US test mean of 10 (SD = 3)</i>	72
Table 9.	<i>Descriptive statistics by gender</i>	78
Table 10.	<i>Correlations between decile/city and test scores</i>	80
Table 11.	<i>Correlation between parent and teacher ratings on the BASC-2</i>	85
Table 12.	<i>Paired samples t-tests for BASC-2 composites across parent and teacher raters</i>	86
Table 13.	<i>Correlation between parent and teacher ratings on the BRIEF</i>	87
Table 14.	<i>Paired samples t-tests for BRIEF composites across parent and teacher raters</i>	88
Table 15.	<i>Correlations of BASC-2 composite and content scales with NEPSY-II scores</i>	91
Table 16.	<i>Correlations of BRIEF composites with NEPSY-II scores</i>	93

TABLE OF FIGURES

<i>Figure 1.</i>	A brief depiction of progress to date for the larger study	36
<i>Figure 2.</i>	Phases of the larger study	38
<i>Figure 3.</i>	BASC-2 means for the total sample, compared to the U.S. norm	69
<i>Figure 4.</i>	BRIEF means for the total sample, compared to the U.S. norm	70
<i>Figure 5.</i>	NEPSY-II means by gender and total sample, compared to the U.S. norm	73

ABBREVIATIONS

Abbreviations are presented in order of appearance within the body of the text

EF	Executive Function
ToM	Theory of Mind
NZ	New Zealand
AR	Affect Recognition
PFC	Prefrontal Cortex
ADHD	Attention-Deficit Hyperactivity Disorder
ASD	Autism Spectrum Disorders
ABI	Acquired Brain Injury
MOH	Ministry of Health
NEPSY-II	Developmental NEuroPSYchological Assessment, Second Edition
BRIEF	Behaviour Rating Inventory of Executive Function
BASC-2	Behaviour Assessment System for Children, Second Edition
TRS	Teacher Rating Scale (BASC-2)
PRS	Parent Rating Scale (BASC-2)
SDH	Structured Developmental History (BASC-2)
CBCL	Child Behaviour Checklist
SDQ	Strengths and Difficulties Questionnaire
WISC-IV	Wechsler Intelligence Scale for Children, Fourth Edition
TEA-Ch	Test of Everyday Attention for Children

CHAPTER ONE

Preface

The intention of this chapter is to provide a brief rationale and overview for the thesis, and to describe the larger study from which the research originated.

On entering the DClInPsych programme, an opportunity arose to work with researchers on a study evaluating whether pesticide exposure is associated with neuropsychological functioning in school aged-children. I was engaged with recruitment and gathering data for the larger study. Ways in which research could be conducted within the context of this larger study were discussed, resulting in permission to use the measures that the larger study were using with their control group (children with little or no exposure to pesticides). Further investigation revealed that the measures selected had yet to be empirically evaluated in New Zealand (NZ).

Information gained from these different assessment tools would allow for evaluation of various measurement issues associated with neuropsychological assessment in middle childhood (age 5 – 12 years) and applied within a NZ cultural context. Executive Function (EF), Theory of Mind (ToM), and Affect Recognition (AR) were of particular interest due to their association with cognitive, social, and emotional development with this age group. An overview of the EF, ToM, and AR literature is reported to provide an outline of the theoretical constructs and assessment issues with this age group.

NZ research to date has provided evidence that performance on neuropsychological tests may not capture cultural differences in performance. Despite these findings, most neuropsychological tests currently used in clinical practice do not have available NZ norms. Cultural differences in reporting, as well as differences between NZ and US schooling systems may also correspond with differences in parent

and teacher ratings of functioning across cultures, although this has yet to be formally assessed. Given the wealth of new research on the development of the prefrontal cortex across childhood, and given potential cultural differences in development, it seemed timely to evaluate assessment of these functions within a NZ cultural context, utilising measures currently used in practice.

A number of subjective and objective assessment tools are being used to assess EF, ToM, and AR in NZ school aged children (without NZ norms) however it is unclear how information gained from these tools relates to each other. Both performance measures (objective) and rating scales (subjective) are utilised in neuropsychological assessment, yet recent studies suggest that subjective and objective assessment tools may not be as closely related as initially thought.

Given the wealth of disparate information on these constructs it is unclear how best to analyse and interpret information gained from assessment tools to inform clinical practice. The current research aimed to clarify how NZ children in middle childhood perform on selected measures of EF, ToM, and AR, how information from these tools relate to each other, and what to make of multi-method assessments in this cultural context.

The aims of the current research are threefold: 1) to evaluate the cultural appropriateness of the selected measures (BASC-2, BRIEF, and NEPSY-II selected subtests) for a NZ non-clinical sample, 2) to evaluate the strength of relationship between parent and teacher reports of functioning, and 3) to evaluate the relationship between subjective (parent and teacher rating scales) and objective (child performance) measures of ToM, AR, and EF.

Very early in data collection, it was observed that there were a number of Critical Items being identified on the BASC-2. As these items potentially indicate

concern, these were discussed in supervision and followed up accordingly. Because this was an unexpected finding during data collection but had potential to be clinically and practically relevant, analysis of these items was included within the first aim of the thesis.

This research adds to the literature in the following ways. First, despite the BASC-2, BRIEF, and NEPSY-II being frequently used in clinical practice with NZ children, there are no previous studies looking at whether there are cultural differences when compared with the normative sample. Second, to this author's knowledge, critical items have not been reported on elsewhere, yet from a clinical perspective this analysis provides useful information in terms of conceptualisation and prevalence of these items within a typically developing group of children. Third, given that EF and ToM are thought to be very closely linked, there has been little consideration of ToM in research that compares parent and teacher reports with child performance.

The following chapter provides an overview of assessment methods and methodological issues in neuropsychological assessment. Chapter Three provides an overview of the current literature on the constructs: EF, ToM, and AR and the likely relationships between them. This chapter then concludes by describing measurement issues that relate to assessment of these domains. Chapter Four provides a brief summary of the gaps in the literature and concludes by presenting the current research. Method describing details related to the collection and analysis of data will be presented in Chapter Five, and preliminary analyses in Chapter Six that pertain to all results.

The results will then be reported across three studies. The first study, described in Chapter Seven, will compare performance from our sample to the norms of each test to determine if these norms are appropriate for use in NZ. The second study, Chapter Eight, evaluates the relationship between parent and teacher reports of functioning and

the third study, Chapter Nine, evaluates the relationship between subjective and objective measures of executive and social functioning. Each study includes a discussion of the results. The final chapter, Chapter Ten, will provide a discussion of the research strengths and limitations, a summary of research findings and a discussion of implications for clinical practice.

Finally, as an additional requirement for completion of the DClinPsych programme, a case study is attached as an appendix to the thesis (Appendix D) which outlines the author's reflections on the research process and the ways in which research has been applied to clinical practice.

CHAPTER TWO

Neuropsychological Assessment

Overview

This chapter will begin by broadly introducing neuropsychology and neuropsychological assessment. It will then discuss current directions in child assessment, particularly the importance of using developmentally and culturally appropriate assessment tools. The chapter will conclude by suggesting that future research is required to evaluate the appropriateness of assessment tools for a New Zealand sample.

Neuropsychological Assessment

Developmental Assessment. The study of neuropsychology is vital to the understanding of brain behaviour relationships (Riccio, Sullivan, & Cohen, 2010) with developmental neuropsychology contributing to our understanding of how these relationships develop over time. Child neuropsychology has only existed as a separate discipline since 1984 (Semrud-Clikeman & Ellison, 2009). Prior to this, children were assessed as if they were miniature adults and many of the formal tests used in these assessments were originally designed for adults (Anderson, Northam, Hendy & Wrennall, 2001; Semrud-Clikeman & Ellison, 2009). Therefore the tests used to identify a child's functioning did not encapsulate the important changes that occur across childhood and adolescence (Semrud-Clikeman & Ellison, 2009). By identifying the normal trajectory of cognitive processes and their behavioural correlates, it is possible to detect early indicators that children may be at risk for experiencing difficulties, and intervene early.

Cultural Differences in Assessment. Assessment tools should not only be developmentally, but culturally appropriate (Carter, Briggs-Gowan & Davis, 2004). NZ is a relatively small country which differs substantially from the cultural context in which most standardised tests are normed. When compared to US children, NZ findings have suggested that children have slower articulation and speaking rates which may contribute to difference in performance on timed tasks (e.g., Robb & Gillon, 2007). There are also documented differences in early schooling experiences, both in structure and curricula (Smith & Elley, 1997) that may contribute to differences in performance and social-emotional functioning when compared to American norms. Even between schools in NZ, cultural differences are noted. McNaughton, Phillips, and MacDonald (2000) suggest that schools with higher rates of Maori and Pacific Island students may have a lower percentage of students meeting educational standards. They also suggest that lower decile schools may be less likely to meet normative standards, as students at these schools may have less exposure to early literacy and language prior to school entry (McNaughton et al., 2000).

Psychometric assessment tools enable clinicians to compare aspects of a client's current functioning to that of the supposedly normative population used to develop standardised norms for the tool. The more the characteristics of the standardisation sample match that of the client, the more valid the assessment tool is at depicting whether the client's functioning is within or outside of the normal range (NZ Psychologists Board, 2013). Ideally then, assessment tools would be developed and normed within the cultural context in which they are to be used. However, most assessment tools currently being used within NZ have not been validated for use with a NZ sample (Barker-Collo, 2003), despite abundant research suggesting that different

cultures may show variations in test performance (NZ Psychologists board, 2013; Ogden & McFarlane-Nathan, 2007; Riccio et al., 2010). For example, Fernando and colleagues (Fernando, Chard, Butcher, & McKay, 2003) found significant differences between the performance of NZ children and the American standardisation sample on the Rey Complex Figure Test. Another group of researchers (Ogden & McFarlane-Nathan, 1997) found substantial differences in the performance of young Maori men on the WAIS-R when compared to American norms.

Cultural differences have also been noted on measures assessing psychopathology. For example, in a group of NZ University students, significant differences were found on the Symptom Checklist-90-Revised (SCL-90-R) when compared to American norms (Barker-Collo, 2003). In the education sector, assessment tools like the Burt Word Reading Test, designed to assess word reading for children aged 6 to 12, were initially adopted in NZ until a national review determined that the words being used were not suitable for the NZ educational context, and the measure was later adapted and revised (Gilmore, Croft & Reid, 1981). These studies suggest that assessment of cultural appropriateness is necessary before simply adopting norms from measures developed overseas.

Assessment Modalities

Numerous neuropsychological assessment tools have been developed and adapted in response to an ever growing literature base on child cognitive development (Peterson & Welsh, 2014; Weyandt et al. 2014). In particular, research on the prefrontal cortex and its association with the development of emotion regulation and behaviour has shaped neuropsychological assessment of these functions in recent years. Both

cognitive performance measures and behaviour rating scales have been developed yet research evaluating relationships between constructs these assessments purport to measure is still catching up.

Assessment of Cognition. Neuropsychological assessment includes the assessment of cognitive, social, and emotional domains of functioning (Riccio et al., 2010). Neuropsychological assessment, both with children and in general, is designed to indicate a person's cognitive strengths and weaknesses and how this may be expressed in their behaviours and daily functioning (Snyder, Nussbaum & Robins, 2006). A typical neuropsychological assessment involves the use of tasks that theoretically correspond to specific cognitive domains or psychological constructs. Thus, performance on such a task is thought to indicate a person's ability in this domain. Neuropsychological assessment also typically involves a clinical interview and the collection of supplemental information (Snyder et al., 2006). The combination of objective assessment tools, clinical interview and supplemental informant reports provides a rounded and valid assessment of the person's abilities (Snyder et al., 2006).

A cognitive assessment provides an objective measure of performance that is less subject to interpretation than other measures (for instance interviews and informant reports using rating scales). However, concerns have been raised as to the applicability of findings from a child's performance on a neuropsychological evaluation to their functioning in real-world settings (Carter et al., 2004). Evidence suggests that performance measures provide a comprehensive assessment of the child's functioning in a particular environment at a given time, but may not necessarily be reflective of the

child's level of functioning outside of that setting (Vriezen & Pigott, 2002) and thus they lack ecological validity¹.

Assessment of Behaviour. Information about a child's social, emotional, and behavioural functioning has typically been gathered through the use of parent, teacher, and self-report ratings of functioning. Parent and teacher reports are useful in providing an assessment of how the child is functioning in areas of adaptive and academic functioning, as well as identifying potential deficits. Rating scales completed by parents and teachers also have the benefit that they may be more predictive of functioning in everyday life than some measures of performance (Anderson, 2002).

Rating scales are quick and efficient, and may provide broader information than that available from an evaluation of performance at one time-point. They may be broadband, examining functioning across a range of areas, or narrowband, examining a specific cognitive domain or aspect of behaviour. They provide a different perspective and also enable clinicians to compare a child's scores to that of a normative sample. For example, a behaviour rating scale may elicit information about a child's typical behaviours, or the frequency with which they engage in those behaviours (from the rater's perspective) which is not necessarily directly observed in session. They also allow for measurement of change throughout therapy. Evidence suggests that parents are accurately able to differentiate between internalising and externalising problems and describe their child's social and emotional strengths (Carter et al., 2004). However, informant reports are typically more subjective in nature than performance measures

¹ The term ecological validity refers to 1.) how representative a task is of contexts or situations in everyday life where the cognitive domain being tested would typically be applied and 2.) how generalizable the findings from the task are, and whether low results are predictive of difficulties outside of the testing environment (Burgess et al., 2006).

and are informed by the rater's relationship and history with the child and the level of concern that the informant attaches to specific behaviours. In this way, different informants may interpret the same behaviour in a different way (Carter et al., 2004).

Due to the level of subjectivity involved, parent reports may be subject to bias (Carter et al., 2004). Reasons for under or over reporting symptoms may include gaining access to services, fear of stigma, or fear of how their parenting skills will be interpreted if their child is seen to be having behavioural problems (Carter et al., 2004). Due to these concerns, multiple informant ratings are recommended as they provide a fuller depiction of the child's functioning across situations and contexts (Banerjee, Watling & Caputi, 2011). Discrepancies are common however, with parents and teachers typically having low to moderate concordance on measures of emotional and behavioural difficulties (Clay, Surgenor & Frampton, 2008; Kennerly, 2013; Achenbach, McConaughy & Howell, 1987) due to children exhibiting different behaviours in different contexts. So too, a child's behaviour may influence parenting behaviour, or the way in which a child is taught in school, ultimately escalating the behaviour in a given context. "Therefore, whether or not the parents reports are actually consistent with the child's behaviour, a child whose parents hold negative appraisals of them and have low expectations – are likely to escalate in problem behaviour" (Carter et al., 2004, p. 113).

To date, both performance measures and rating scales have been used in practice with the understanding that both have their limitations. Less clear, is how to interpret multi-method assessments of the same theoretical constructs.

Summary

Neuropsychology is the study of brain-behaviour relationships.

Neuropsychological assessments should be both developmentally and culturally appropriate for the population being studied. In clinical practice, both performance measures and rating scales are utilised. Performance measures are often used to infer behaviour, despite evidence that they may not be as highly related as first thought. In order to determine how information produced by current assessment tools (e.g., multi-method assessments) can inform clinical practice with NZ children, it is important to evaluate these tools and the relationship between them within our unique cultural context.

CHAPTER THREE

Executive Functioning, Theory of Mind and Affect Recognition

Overview

As indicated in Chapter Two, assessment of executive function (EF) and the role of the prefrontal cortex (PFC) in the regulation of emotion and behaviour has received considerable research interest over the last 30 years (Szczepanski & Knight, 2014; Hale & Fitzer, 2015). Given its association with the development of cognition, academic advancement, and the ability to form and maintain social relationships (Anderson, 2002), this is not surprising. Theory of Mind (ToM) and Affect Recognition (AR) have also been associated with the PFC and developments in cognitive, social, and emotional development.

Neuropsychological assessments are often required after a traumatic brain injury, which children are likely to sustain in falls or collisions (e.g., Feigen et al., 2013). These are frequently sustained head first and are likely to damage the frontal regions of the brain where EF and ToM processes occur (e.g., Verger et al., 2000). Hence, accurate assessment of these processes throughout development is vital.

This chapter will first introduce the constructs EF, ToM, and AR, and the history around how they developed. The relationship between these constructs will be discussed, and the chapter will conclude with the suggestion that further assessment of these functions in middle childhood, and with a NZ sample, is indicated.

Executive Function

Executive Function (EF), little studied until the 1970s, is now considered an essential domain of cognitive functioning (Goldstein, Naglieri, Princiotta & Otero, 2014). In essence, EF involves numerous processes related to purposeful, goal-directed action, though this definition is constantly evolving. Wasserman and Wasserman (2013) estimated that 3000 scientific articles published over 15 years had yet to reach a consensus on exactly what EF is. In fact, in a recent review of the literature, Goldstein and colleagues (Goldstein et al., 2014) listed 33 different definitions that have been used to describe EF. Despite this, researchers in the field agree that executive processes are associated with the PFC and “play an important role in a child’s cognitive functioning, behaviour, emotional control, and social interaction” (Anderson, 2002, p. 71).

EF includes such processes as initiation and planning, the ability to store and manipulate information in working memory, self-regulatory behaviours, and the ability to problem-solve, select and execute an action and then monitor and respond to feedback (McCloskey & Perkins, 2013). Best, Miller, and Jones (2009; see also Garon, Bryson & Smith, 2008) report a general convergence in the recent literature toward viewing EF as a collection of “related but separable components” (p. 183, Best et al. 2009), namely focused on Miyake et al.’s (2000) theory that EF consists of three key elements: inhibition, shifting, and updating of working memory. In general, research suggests that children begin to develop EF skills in preschool and that these skills continue to develop through to adulthood. Inhibition performance appears to be particularly important during the preschool years when children might be more distractible, however significant improvements have also been noted later in childhood, particularly between ages 5-8 (Romine & Reynolds, 2005). Best and colleagues (2009) acknowledge that there is some evidence to suggest linear development in working

memory abilities, however age-related changes in performance vary as a function of task complexity and scoring method. In terms of shifting ability, steady improvements were reported across childhood and adolescence. In terms of planning, children are thought to be able to plan up to three moves by middle childhood and develop more complex plans in adolescence. For further review, see Best et al (2009).

EF processes are strongly related to social understanding and behaviour, but can also be associated with emotional reactions (Gioia, Isquith, Guy & Kenworthy, 2000b; Anderson, 2002). Children with EF impairments may experience disruptions in mood, energy level and initiative, as well as other emotional and behavioural difficulties such as poor impulse control, perseveration difficulties or problems with monitoring or regulating behaviour (Anderson, 2002).

Deficits in EF have also been associated with an increase in behavioural problems (Beauchamp & Anderson, 2010). Anderson (2002) notes that children with EF deficits may lack insight into the social appropriateness of their behaviours and may have difficulty adhering to social norms. He further points out they may also have difficulty with nuances of conversation, for example humour or sarcasm (Liddle & Neddle, 2006; Hughes & Leekam, 2004). They may be unable to adapt and learn from previous mistakes and as a result of these impairments, children with EF difficulties may have difficulty forming and retaining positive relationships. It is important to note however, that there are substantial individual differences in EF development and presentation.

More recent research is beginning to tease apart cognitive and affective aspects of EF. An assumption of traditional neuropsychological assessment was that EFs are domain-general in that they function in the same way across various contexts (Hongwanishkul et al., 2010). However, some authors suggest that unlike the original

domain-general view of EFs, different processes may be elicited by different contexts (e.g., Hughes & Graham, 2002; Zelazo & Muller, 2002). This theory suggests that there are *hot* EFs which become active “in situations which require affect regulation and motivation” (Hongwanishkul et al., 2010, p. 618) and *cool* EFs which are elicited by more cognitive conceptual tasks (Hongwanishkul et al., 2010). *Hot* EFs have been linked to the ventro medial PFC where *cool* EFs have been linked to the dorsolateral PFC (e.g., Hongwanishkul et al., 2010; Hale & Fitzer, 2015). Researchers in this field tend to report that ToM is a subdomain of hot EF (e.g., Burgess et al., 2006).

EFs are thought to be responsible for cueing and directing other cognitive functions (McCloskey & Perkins, 2013). As a result, there is a degree of overlap and interrelationships with other areas of psychological and neuropsychological function. While theoretical understanding of EF has been derived from knowledge of brain – behaviour relationships, other psychological constructs (when initially conceptualised) such as ToM, did not derive theoretical underpinnings from knowledge regarding brain function. “As a result, these psychological constructs all have various elements of the multidimensional construct of executive function deeply embedded in their core theoretical conceptions and operational definitions with little or no explicit acknowledgement of this fact” (McCloskey & Perkins, 2013, p. 45).

Theory of Mind (ToM).

Relative to EF, ToM is a newer area of research that is only now beginning to be understood. While EF relates to a broad range of cognitive functions, ToM is a more narrow construct that relates specifically to social functioning. It involves attribution of mental states, such as belief, to self and others (Austin, Groppe & Elsner, 2014), and the ability to infer intentions and behaviours from another’s thoughts, beliefs, and emotions

(Bosacki & Wilde Astington, 1999). Introduced by Premack and Woodruff in 1978, it has since become one of the most extensively researched topics in developmental psychology (Saracho, 2014; Wellman, Cross, & Watson, 2001). Like EF, it has been associated with the development of empathy, academic success, and the development of peer, teacher, and family relationships, which are vital for long term psychological functioning (Crowe, Beauchamp, Catroppa & Anderson, 2011; Banerjee et al., 2011), and is described as a fundamental skill which most people develop in infancy and early childhood (Berk, 2006). As with EF, there are also substantial discrepancies across studies in the way ToM is defined, and in the research methods and assessment tools used, which has led to substantial confusion around how best to conceptualise and assess ToM abilities (Schaafsma, Pfaff, Spunt & Adolphs, 2015).

Within neuropsychology research, ToM development has been studied within the context of clinical disorders in which ToM development has been found to be impaired or delayed, such as Schizophrenia, Attention-Deficit Hyperactivity Disorder (ADHD), Autism Spectrum Disorders (ASD), and following Acquired Brain Injury (ABI). Research with typically developing children is in its infancy (Caputi, Lecce, Pagnin & Banerjee, 2012). Recent findings also indicate that increased ToM is related to an increase in prosocial behaviours (Slaughter, Dennis & Pritchard, 2002; Caputi et al., 2012) which suggests that intervention targeting ToM development could mediate problem behaviours.

Typically studied prior to entry into formal schooling, the majority of research in this area focuses on first order ToM, which is the ability to understand another's mental states (Bosco, Gabbatore, & Tirassa, 2014). First order ToM is assessed using tasks such as the False-Belief task (Baron-Cohen, Leslie & Frith, 1985; Gopnik & Astington, 1988; Wimmer & Perner, 1983) where to be successful, children need to differentiate between

what they know and what another person thinks, and demonstrate understanding that the other's belief is false. Success on these tasks is typically achieved by age four (Wellman et al., 2001; Austin et al., 2014). Second order ToM involves the ability to reason about what one person thinks of another's thoughts, and is typically developed by age six (Baron-Cohen & Swettenham, 1997).

Individual differences in ToM development have been reported both in clinical and typically developing populations (Hughes, 2011; Caputi et al., 2012; Banerjee et al., 2011; Liddle & Nettle, 2006). Recent evidence suggests that ToM development does not stop at six years as previously thought, but instead develops gradually throughout childhood (Hughes, 2011), with recent findings suggesting that middle childhood may be “a particularly important period for the manifestation of sociocognitive skills in real-life peer contexts” (Caputi et al., 2012, p. 266). However, with the exception of a few (e.g., Devine & Hughes, 2013; Im-bolter, Agostino & Owens-Jaffray, 2016; Apperly, Samson & Humphreys, 2009; Bosacki & Astington, 1999), there has been very little research on ToM assessment with children over six years.

While originally rooted in developmental psychology, findings from neuroscience and neuropsychological research are now being applied to ToM (Mahy, Moses & Pfeiffer, 2014; Schaafsma et al., 2015). Like EF, ToM has more recently been conceptualised as two distinct constructs, affective and cognitive ToM (Sebastian et al., 2012), and as with the concept of *hot* and *cool* EFs, this theory is supported by findings from a neuropsychological study which demonstrated that people with lesions to the ventromedial PFC exhibited specific impairments on affective reasoning tasks (Shamay-Tsoory, Tibi-Elhanany, & Aharon-Peretz, 2006). Affective ToM is thought to involve emotional reasoning or empathy and develops later than cognitive ToM, which requires reasoning about thoughts and beliefs (Shamay-Tsoory et al., 2006).

When evaluating a child's social functioning, it is important to consider how the child *expresses* their emotions and behaviours, and not simply their *understanding* of emotions and behaviours in others. The association between cognitive ToM and behavioural functioning, for example, is only beginning to be evaluated despite it being largely accepted that children with a developed ToM will behave in more socially appropriate ways than those without (Banerjee et al., 2011).

Research to date suggests that the way in which children understand emotions and behaviours may not necessarily be reflected in the way they express them (Hughes, 2011). In fact, just because a child has a developed ToM does not necessarily mean that they will use that ability in everyday situations (Hughes, 2011). This suggests that performance on cognitive tasks may differ from information provided by informant report. Caputi and colleagues (2012) postulate that a higher degree of ToM is required in order for children to utilise these skills in socially positive ways, which is consistent with findings suggesting middle childhood may be a more appropriate time to assess ToM understanding and its behavioural correlates.

Affect Recognition (AR).

AR is another psychological construct, typically defined as the ability to recognise emotions in the facial expressions of others. As with ToM development, the ability to recognise affect typically develops prior to school entry. Findings are mixed however, as to whether AR continues to develop throughout childhood. Some authors write that children will be able to perform at adult level on these tasks by approximately age 5 (e.g., Wong, Beidel, Sarver & Sims, 2012), where other studies report that AR continues to develop across childhood (e.g., Demopoulos, Hopkins & Davis, 2013).

Deficits in AR have been reported in both child and adult populations in a number of clinical disorders, including Bipolar Disorder (McClure, Pope, Hoberman, Pine & Leibenluft, 2003), ASD (Harms, Martin, & Wallace, 2010), ADHD (Demopoulos et al., 2013), Social Phobia (Simonian, Beidel, Turner, Berkes, & Long, 2001), Depression (Lenti, Giacobbe & Pegna, 2000), Schizophrenia (Brune, 2005; Yang et al., 2015), and following Traumatic Brain Injury (Henry, Phillips, Crawford, Ietswaart, & Summers, 2006). Early studies have begun to show that training programs targeting AR may improve emotion recognition for people with Schizophrenia (e.g., Silver, Goodman, Knoll, & Isakov, 2004; Wolwer et al., 2005).

ToM and AR are often simultaneously affected but have been found to represent distinct abilities (Brune, 2005). The brain regions associated with AR are the amygdala and the orbitofrontal cortex (Frith & Frith, 1999). Like ToM, gender differences in the development of the amygdala imply that there may also be individual differences in performance on tasks of AR (Blakemore & Choudhury, 2006).

Relationship between Constructs

EF and ToM. For several years researchers have been identifying similarities particularly in the development of EF and ToM (Carlson, Moses & Breton, 2002). First, EF and ToM skills both develop prior to school entry, and continue to develop through to adulthood (Carlson et al., 2002; Anderson, 2002). Second, the prefrontal region is implicated in the development of both EF and ToM, which is thought to indicate commonalities in cognitive processing (Baron-Cohen & Swettenham, 1997; Carlson et al., 2002; Anderson, 2002). Third, many individuals who exhibit deficits in ToM similarly exhibit deficits in EF (Carlson et al., 2002). Fourth, like EF, ToM deficits have also been described in ADHD, ASD, and following ABI (Anderson, 2002). Children

with EF difficulties may present similarly to children with ToM deficits (Anderson, 2002), and as with ToM development, there are substantial differences in EF abilities across individuals.

As previously stated, similar conceptualisations of *hot* and *cool* EFs and affective and cognitive ToM have been proposed. In a study assessing the relationship between EF and affective ToM in adolescence, Vetter, Altgassen, Phillips, Mahy & Kliegel (2013) found that inhibition was the single most important predictor of affective ToM, though other EFs were also correlated.

Although there has been an abundance of recent research, no cohesive argument for how EF and ToM relate in middle childhood has emerged. In fact, many of these theories contradict each other (Perner & Lang, 1999). The major point of discrepancy is whether or not EF and ToM are distinct constructs or whether ToM is a subset within the larger construct of EF (e.g., Zelazo & Muller, 2002).

Proponents of the expression account (e.g., Perner & Lang, 1999) posit that EF is required for the expression of ToM understanding and that EF skills are “essential in guiding appropriate adjustment of behaviour” (Tonks, Yates, Williams, Frampton & Slater, 2014, p. 154). This makes sense given that EF is thought to be involved in metacognition and controlling other cognitive processes.

Carlson and colleagues (2002) proposed that EF is employed in skill learning during development, and that the strong relationship between EF and ToM in preschool is due to children having developed EF skills. The underlying assumption being that for a child to be self-aware, and therefore understand that they are distinct from other people, requires self-monitoring, a skill typically associated with EF (Perner & Lang, 1999). Similarly, more recent findings postulate that a child’s development of ToM is dependent on their level of EF (Carlson, Koenig & Harms, 2013).

Alternatively, other research has suggested that the development of ToM skills precedes the development of EF. This theory posits that with the development of ToM comes an increased understanding of mental functions, and this understanding then leads the child to develop control over their mental functions and actions (Perner & Lang, 1999). Researchers in this area believe that a “common mental factor” underlies performance on both ToM and EF tasks (Perner & Lang, 1999, p. 339), and this model is thought to explain the combined EF and ToM deficit profile in both ASD and Schizophrenia (Perner & Lang, 1999).

These theories suggest that EF and ToM are functionally dependent (Perner & Lang, 1999), and begin to explain why these domains may often be simultaneously affected in certain symptom profiles. However, they fail to account for findings that they are not *always* simultaneously affected. For instance, if EF precedes ToM this discounts the possibility that someone may have EF difficulties and still have an intact ToM. Yet Tager-Flusberg and colleagues reported that children with Williams or Prader-Willi Syndromes may perform poorly on measures of EF and yet be successful on a ToM false-belief task (Tager-Flusberg, Sullivan & Boshart, 1997). These theories also fail to describe why children tend to become proficient on ToM false belief tasks and EF inhibition tasks (typically associated with ToM development) at the same age (Perner & Lang, 1999).

Based on their own findings and that of their meta-analysis, Perner and Lang (1999) determined that during pre-school development both processes are mutually dependent on each other. In their paper they concluded that it was still unclear to what extent this relationship was influenced by brain structures and their degree of maturation at this stage of development. This suggests that the relationship may differ at different

stages of development. The relationship between EF and ToM in middle childhood, for instance, is less clear.

EF and AR. Recent research has also begun to investigate the relations between AR and EF. A number of studies have found correlations between AR and EF skills both in typically developing and clinical populations (e.g., David, Soeiro-de-souza, Moreno, & Bio, 2014; Kohler, Walker, Martin, Healey, & Moberg, 2010), though no studies were found that evaluated this relationship in children.

In the *Developmental NEuroPSYchological Assessment, Second Edition* (NEPSY-II; Korkman, Kirk & Kemp, 2007), intercorrelations between the AR subtest and the EF subtests revealed little to no relation between the measures (e.g., Animal sorting: $r = .11$; Inhibition $r = .19$; Switching $r = .14$). In a similar way, a study evaluating AR, ToM, and EF in adults with Schizophrenia versus healthy controls, found a strong relationship between ToM and EF, but AR was not related (Brüne, 2005).

ToM and AR. Recently, a new Social Perception domain was added to the NEPSY-II, which consisted of two subtests, ToM and AR. No rationale was provided for combining AR and ToM other than to aid assessment of ASD. The authors report that the intercorrelations between NEPSY-II subtests “provides strong evidence that the structure of the NEPSY-II is sound” (Korkman et al., 2007, p. 81) when in fact the intercorrelation between the two Social Perception subtests is only $r = .21$ (Korkman et al., 2007). ToM was more strongly correlated with language and memory NEPSY-II subtests, such as Comprehension of Instructions ($r = .39$), Narrative Memory ($r = .41$), and Sentence Repetition ($r = .37$). It seems that while the NEPSY-II has chosen to link

ToM with AR due to their common co-occurrence in ASD, the two domains are not strongly related, and in fact it seems that ToM appears to be more strongly related to measures of EF and Language.

In sum, the prefrontal cortex is associated with the development of three important domains of functioning: EF, ToM and AR. These functions are thought to develop across middle childhood yet studies involving this population are limited. The picture so far suggests that ToM seems to be strongly related to EF, to the point that many consider it to *be* an executive process. These concepts were originally developed in isolation, yet over time ToM has become superimposed on EF. The two concepts are not a perfect fit however, and this leads to confusion over how they should best be understood and assessed. AR on the other hand, does not appear to be strongly related to either ToM or EF. Although AR and ToM appear to be two different concepts, they have been grouped into one Social Perception domain in the NEPSY-II. Due to the blurred boundaries between how EF and Social Perception are defined and conceptualised, naturally there is likely to be some overlap in the way in which these domains are assessed as well.

Measurement Issues

Cultural appropriateness of EF, ToM, and AR measures. As described in Chapter One, frequently used neuropsychological assessment tools have not been evaluated for use with NZ children, despite evidence to suggest that performance and ratings may differ across cultures.

Studies from outside of NZ have identified that neurocognitive functions, including social perception, may vary across cultures (Rosenqvist et al., 2017). In a

study looking at cultural differences in performance on the NEPSY-II, Rosenqvist and colleagues (2017) found differences on a range of subtests when comparing performance for children from three different Western countries (Finland, Italy and the United States). While they and others (e.g., Schaafsma et al., 2015) identified differences in ToM performance across cultures, others have reported no difference (e.g., Sabbagh, Xu, Carlson, Moses & Lee, 2006). Wang and colleagues (Wang, Devine, Wong & Hughes, 2016) suggest that pedagogical differences may have more bearing on ToM performance than cross-cultural differences. Shahaieian et al. (2014) suggest that there are differences across cultures in patterns of development, for example while there was no difference in overall ToM scores, more Iranian children showed an understanding of sarcasm, whereas more Australian children demonstrated awareness of belief diversity. In contrast Wellman et al. (2006) found no differences in the sequence of ToM development in a study comparing results of Chinese children in Beijing with English speaking children from America and Australia. Reports of EF differences cross-culturally also differ across studies (e.g., Wang et al., 2016).

Individual differences in performance. For the most part, girls tend to perform higher than boys on tasks of social understanding (eg., Bosacki & Wilde Astington, 1999). However this is not always the case (e.g., Liddle & Nettle, 2006). Findings suggest that in both pre-school (e.g., Bosacki & Wilde Astington, 1999) and adult populations (e.g., Baron-Cohen & Wheelwright, 2004), girls perform better on ToM tasks. In a study with school-age children, Liddle and Nettle (2006) found that boys performed better, although these results have yet to be replicated.

Gender differences in ToM skill development may occur as a result of differences in rate of brain maturation. Indeed, gender differences in brain development have been indicated in ASD, ADHD, and schizophrenia (e.g., De Bellis et al., 2001; Lai et al., 2015; Arnett et al., 2015), all of which are associated with deficits in ToM and EF. EF skills however, are thought to develop similarly for both boys and girls (Anderson, 2002).

Evidence suggests that ToM performance also varies as a function of socioeconomic status, with children from more privileged schools performing better on ToM tasks (Liddle & Nettle, 2006), though this study was preliminary and did not control for EF or general intelligence. There are also thought to be urban/rural distinctions in EF development. For example, pesticide exposure, more common in rural populations, can affect neuropsychological functioning in children (Roberts & Carr, 2012). Due to these differences, it appears that gender and socioeconomic status are important to consider when evaluating social and EF development.

Developmental EF assessment. EF assessment is particularly difficult during early development, as the typical development of EF skills is yet to be conclusively established (Wasserman & Wasserman, 2013). Typical functioning is important to determine, as in order to evaluate whether a child differs from the norm, a norm first needs to be established. While numerous studies have attempted to establish typical development of EF, there has been very little consistency in how it has been assessed, leading Wasserman and Wasserman (2013) to conclude that “the study of EF in children is composed of a patchwork of skills and populations with some assessments focusing on neuropsychological correlates and others focusing on manifest behaviour” (p. 88).

Assessment of EF has typically been conducted within the context of a neuropsychological evaluation as findings suggest that EF assessment should encompass tasks that are “novel, complex, and involve the integration of information” (Anderson, 2002, p. 74). However, evidence suggests that children may behave differently in novel situations, which limits generalisability of findings from an assessment of performance. If it is the case that results of a neuropsychological assessment have limited predictive validity of real-world behaviours, then additional information is required (Anderson, 2002). This additional information is typically collected via rating scales such as the Behaviour Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy & Kenworthy, 2000a) which may be completed by parents or teachers to collect data on how the child functions in the home and school environments.

Developmental ToM Assessment. As previously mentioned, one of the most heavily researched assessment tools for ToM function is the False Belief task, commonly used in preschool assessments. The ecological validity of the False-Belief task has been criticised as it is unclear to what extent it reflects the child’s use of ToM in real life settings (Caputi et al., 2012). Tasks assessing first and second order ToM are also limited in that they have early ceiling effects (Banerjee et al., 2011). While these tasks are useful for determining whether the child has the ToM of a typically developing five or six year old they may not be developmentally appropriate for assessing these skills in older children (Banerjee et al., 2011). Shamay-Tsoory and colleagues (2006) have also reported that most traditional ToM tasks have assessed cognitive ToM, and emotional reasoning was often not considered.

Recently, a ToM subtest was included in a new Social Perception domain in the NEPSY-II, allowing for assessment across childhood (ages 3 – 16). The Social Perception domain consists of two subtests, ToM and AR. The Social Perception domain was introduced in response to research highlighting an increase in Autism Spectrum Disorders (ASD; Kemp 2007). As a deficit in Social Perception, particularly ToM, is characteristic of ASD, this domain was designed to aid assessment of this disorder (Korkman et al., 2007). The addition of these measures into the NEPSY-II allows for EF, ToM, and AR to be assessed with one tool based on a single normative sample as opposed to selecting a range of different instruments.

The Behaviour Assessment System for Children, Second Edition (BASC-2; Reynolds & Kamphaus, 2004) is a broadband behavioural assessment measure which now allows for the assessment of executive behaviour as well as a broad range of other clinical issues, including assessment of adaptive skills. Given that ToM abilities are associated with an increase in prosocial behaviour (Banerjee et al., 2011), rating scales of adaptive skills may provide insight into ToM abilities.

Overlap in EF and ToM assessment. Early research has suggested that the relationship between EF and ToM may be an artefact of ToM assessment tools also requiring inhibition or other processes typically associated with EF (Carlson et al., 2002). In research conducted with preschool age children for instance, EF tasks were found to require the understanding of mental states to inhibit previous responses and ToM tasks required self-monitoring and behavioural regulation processes commonly associated with EF (Perner & Lang, 1999). Some researchers go so far as to say that there are no cognitive assessment tools available that do not in some way involve EF skills (e.g., Anderson, 2002), though these reports were made prior to the introduction

of the NEPSY-II. The expansion of the NEPSY-II to include a Social Perception domain allows for direct comparisons in performance on these domains from preschool to adolescence (Kemp, 2007).

Unfortunately, despite the authors of the NEPSY-II manual acknowledging that “children with significant social impairments will usually have concurrent deficits in executive functioning” (Korkman et al., 2007, p. 18), when reporting correlations with the Delis-Kaplan Executive Function System (D-KEFS) the ToM subtest was not included. Intercorrelations are also absent between ToM and the Inhibition-Switching, and Animal Sorting subtests of the EF domain. Due to the extensive overlap previously described between these two cognitive domains, it is unclear whether the NEPSY-II ToM task measures ToM alone, or simply EF ability. Thus evaluating comparisons with behavioural correlates may provide an interesting avenue for further research.

Subjective versus objective assessment. The correlation between the BRIEF and performance measures of EF is low-moderate, which suggests that the information provided by these measures does not completely overlap (Anderson, 2002; Toplak, Bucciarelli, Jain & Tannock, 2008). While some researchers propose that subjective reports and performance based measures provide complementary information (Isquith, Roth & Gioia, 2013), others have argued that they assess different constructs (Rubin & Pepler, 2013). Isquith and colleagues (2013) describe a number of studies, including their own, which found either minimal or no significant correlations between rating scales and performance measures of EF. These authors suggested that the reason there may be minimal correlation between the two is that the performance measures typically selected for these analyses have assessed only *cool* EFs, where rating scales such as the BRIEF assess *hot* EFs. Thorough review of the relevant literature revealed no research

to date that has examined the link between subjective informant ratings of a child's emotion and behaviour and their objective performance on neuropsychological measures of Social Perception in middle childhood.

Summary

Given differences noted in performance across cultures, determining how NZ children perform on these measures is important to consider to inform clinical decision-making. Objective and subjective measures are both utilised in practice to reduce problems of ecological validity and response bias, yet it is unclear whether they yield similar or differing information. Research to date suggests that the discrepancy between objective performance on traditional neuropsychological assessments and subjective rating scales may be a result of performance measures targeting cognitive components and rating scales assessing affective components of these constructs (Isquith et al., 2013). Now that the NEPSY-II includes assessment of affective components (ToM and AR) the relationship between subjective and objective measures can be more effectively studied.

CHAPTER FOUR

The Proposed Research

Gaps in the Current Literature

As stated in the previous sections, there have been significant advances in the field of paediatric neuropsychology, with a number of assessments being developed to assess executive and social functioning in children. Developments in EF, ToM and AR research have led to remarkable similarities in the way they are conceptualised and assessed (Carlson et al., 2002), with all three now being assessed simultaneously within the same neuropsychological assessment tool, the NEPSY-II.

Subjective (broadband and narrowband) and objective measures have been developed, yet given the level of overlap between the constructs and assessment tools combined with the lack of NZ norms, it is unclear how best to interpret information gained from these measures for NZ children. Previous studies indicate that subjective and objective measures may not be related, yet these studies have not considered affective components such as ToM and AR. Now that the NEPSY-II allows for assessment of affective tasks as well as *cool* EF tasks, it is possible to evaluate this relationship further.

Few studies have evaluated assessment of EF, ToM, and AR in middle childhood (between preschool and adolescence), despite evidence that suggests that crucial developments may occur during this stage of development (Caputi et. al., 2012). Intuitively, middle childhood also seems an important time to assess social functioning as it is at this stage that children have typically developed the pre-requisite skills to form social bonds and act in socially appropriate ways. For those who have not yet developed these skills, intervention using training programs that target these skills may be successful (Dodell-Feder, DeLisi & Hooker, 2014). Early identification and

intervention may prevent social and emotional problems progressing into chronic difficulties (Carter et al., 2004). Due to the rapid, non-linear development of executive skills during childhood (Anderson, 2002), understanding how assessment of these skills relate to ToM and AR at various stages may improve accuracy of assessment and intervention of children with behavioural difficulties.

Finally, no studies to date have evaluated the cultural appropriateness of the BASC-2, the BRIEF, or the NEPSY-II for NZ children, though they are utilised in both research and in clinical practice. Of particular interest are the BASC-2 critical items. These items are identified when scoring the BASC-2 as specific problem behaviours that may be of clinical importance (Reynolds & Kamphaus, 2004). Though the prevalence of these items has not been recorded even in the United States where this test was developed and standardised, it may be of clinical importance to determine how frequently these behaviours occur within a NZ population, particularly if these behaviours are thought to be indicative of need for concern.

Aims

The aim of the current thesis is to evaluate the extent to which subjective parent and teacher reports on the BASC-2 and BRIEF are associated with child performance on the NEPSY-II Social Perception and EF subtests, and determine whether these tools are suitable for use in the NZ cultural context. The results will be reported in three separate studies to address the following aims.

1. *To evaluate the cultural appropriateness of the BASC-2, the BRIEF, and the NEPSY-II subtests for a NZ sample*². This study will comprise of three parts.

² Included a wide range of cultures. Ethnic distribution of sample discussed on pages 39 and 40.

Part One. Due to the fact that the BASC-2 Critical Items may indicate cause for concern, these scores will be analysed first. BASC-2 critical items have not been reported in previous research, however it is expected that there will be a relatively low prevalence of these items within our (non-clinical) sample.

Part Two. Part Two will compare performance of the NZ sample on the BASC-2, BRIEF, and NEPSY-II measures to American norms where the tests were developed. It is expected that there will be differences in performance and reporting between our sample and the US standardisation sample.

Part Three. Part Three involves an evaluation of other factors that are thought to cause individual differences in performance including gender and socioeconomic status. As differences have been found in previous research, it is hypothesised there will also be differences in our sample.

2. *To evaluate the strength of relationship between parent & teacher reports of functioning.*

This study examines the extent to which parent and teacher ratings on the BASC-2 and BRIEF are related. Previous US findings have reported low to moderate concordance between parent and teacher ratings of functioning. It is hypothesised that similar results would be found in the current study.

3. *To evaluate the relationship between subjective and objective measures of Social Perception and EF.*

This study examines whether parents or teachers ratings of child behaviour (BASC-2 and BRIEF) are associated with the child's performance on ToM, AR, and EF tasks (NEPSY-II). Evidence to date is mixed on the relationship between subjective (e.g. rating scales) and objective (performance) measures, where some researchers posit that they provide complementary information and others suggest they assess different things. As there is no clear indication of whether these measures are related, an exploratory investigation will be undertaken in the present study.

Significance of Research

So far, very few research studies have addressed issues relating to assessment of EF and ToM in middle childhood, despite it being noted that important changes may occur during this time. The proposed research will add to the literature base in several ways. First, reporting prevalence of BASC-2 Critical Items is a novel contribution of this thesis. Second, this research will provide an evaluation of whether the BASC-2, BRIEF, and NEPSY-II test norms are appropriate for administration with NZ children. Third, this thesis will contribute an in-depth analysis of the relationship between objective performance on EF and Social Perception tasks with both broadband and narrowband informant ratings of behaviour from parents and teachers. And finally, this thesis will discuss what these findings mean and how they inform clinical assessment and intervention for this age group.

Research Context

The research was conducted in the context of a larger ongoing study which is evaluating the relationship between pesticide exposure and neuropsychological functioning. To avoid confusion, any reference to the current research refers to that conducted for this thesis. The larger study will be referred to as such.

For the larger study, children were recruited from areas that fall into one of three pesticide exposure groups (low, medium, or high) so as to assess risk of neuropsychological deficits across levels of exposure. The current research is based on analysis of data from the low exposure group which was collected first.

To evaluate general cognitive functioning, the larger study used selected subtests from the NEPSY-II, the Wechsler Intelligence Scale for Children, 4th Edition (WISC-IV) and the Test of Everyday Attention for Children (TEA-Ch) supplemented by teacher and parent responses on a general questionnaire, the BASC-2 (which provides an indication of the child's level of emotional and behavioural functioning across the home and school contexts) and the BRIEF (which provides an indication of the child's EF ability).

Prior to the evaluation of the effect of pesticide exposure on neuropsychological functioning in New Zealand children, it was important to determine the extent to which the measures of neuropsychological functioning being used were appropriate for a typical NZ sample. Further, it was also important to determine how these measures relate to each other and whether parent or teacher reports of behaviour provide additional useful information. In this way, the current research provides the initial data analysis for the larger study to then assess the appropriateness of the measures before comparing performance across levels of exposure.

The author's contribution to the larger study is depicted in Figure 2. Note that the author was responsible for scoring of questionnaire data, and was one of three researchers involved in administering and scoring neuropsychological assessments. The rest was done in collaboration with researchers from the Centre of Public Health Research and the School of Psychology at Massey University. The author was involved in brief discussions about the project in 2013, but did not draft the ethics or funding applications.

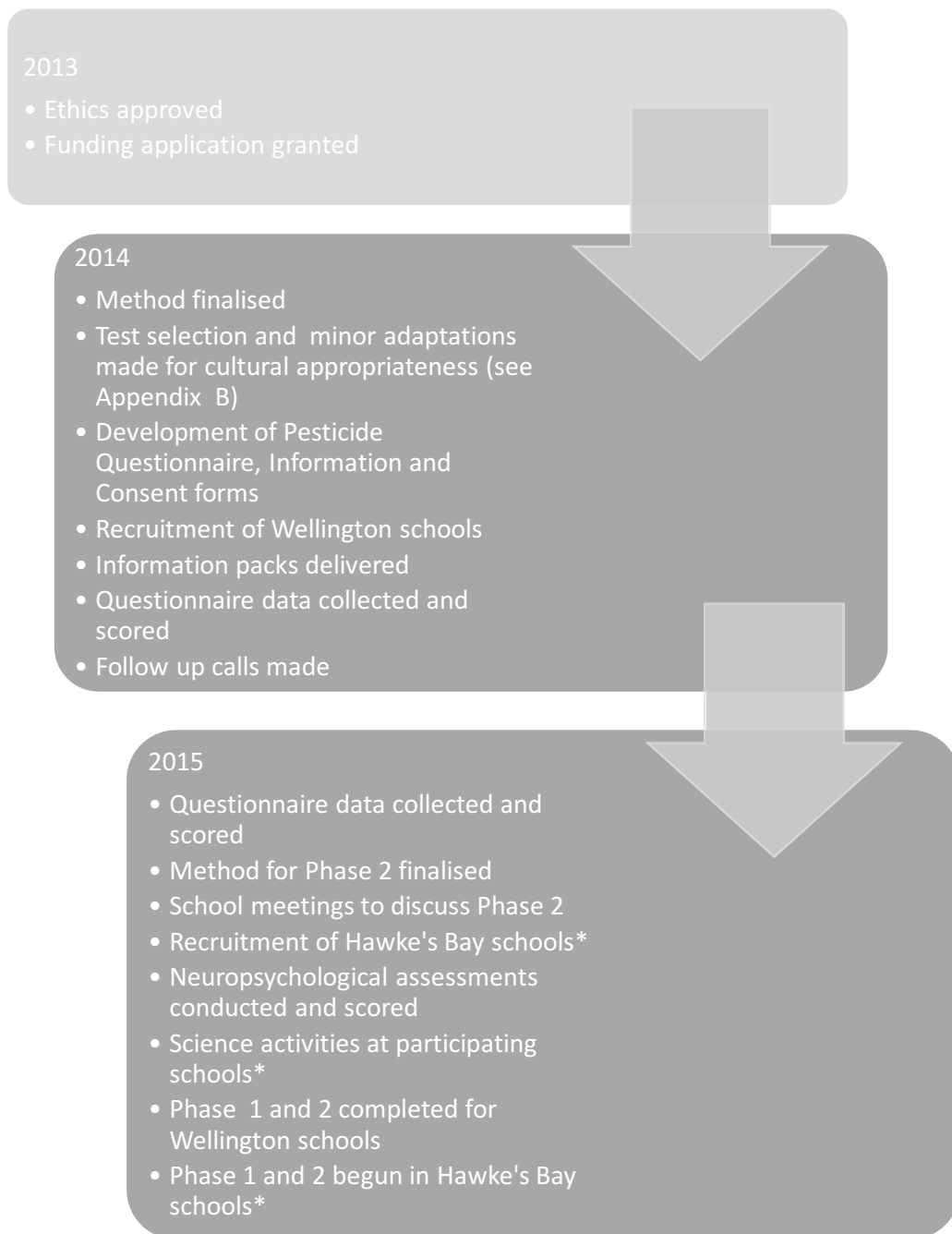


Figure 1. A brief depiction of progress to date for the larger study.

* The author was involved in all aspects of the study in collaboration with the research team for the years 2014 and 2015, except for recruitment and data collection in Hawke's Bay and in the science activities that were held at participating schools to thank them for their participation. In 2016, the author was completing her internship and was not working with the larger study.

CHAPTER FIVE

Method

Prior to recruitment, ethical approval was obtained for the larger study from the Central Health and Disability Ethics Committees (13/CEN/134). As no new measures or methodology were introduced, that ethical approval also covered the current research. The current study involved collaboration with the schools, students, and parents from 11 schools in Wellington and Hawke's Bay, and researchers from both the School of Psychology and Centre of Public Health Research (CPHR) at Massey University.

Procedure

The larger study was conducted in three phases, as depicted in Figure 2. Only Phase 1 and 2 are reported in this thesis, as Phase 3 involved the collection and analysis of urine and dust samples to evaluate Pesticide exposure, which were not relevant for the current research. The current research then, is a cross-sectional investigation with data gathered during two phases; the first consisted of brief self-report questionnaires completed by the parent of each participant, the second consisting of teacher versions of these same questionnaires completed for each participant and neuropsychological evaluation of the child's performance on EF and Social Perception tasks.

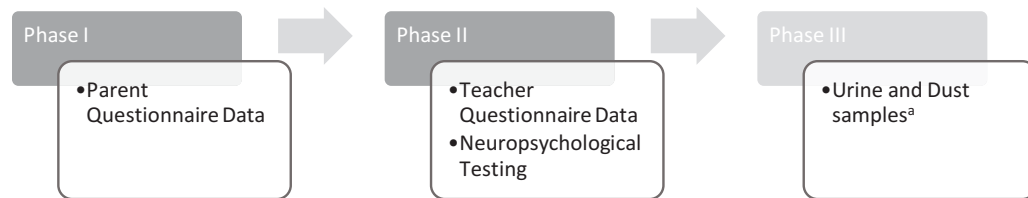


Figure 2. Phases of the larger study.

Note. Phase I, II and III are completed across all three exposure groups.

^aThe author was not involved in Phase 3 of the study.

Phase 1. Phase 1 data collection took place in October-November 2014 and April-May 2015. Phase 1 of the study involved gathering information of the child's functioning within the home. Across the 11 schools, 3003 parents were sent home a pack which included information and consent forms. This could be returned by post or via the school and collected by the research team.

Upon return of the consent forms, questionnaires were sent by post to the parent. Questionnaires included the BASC-2 and BRIEF parent forms as well as a questionnaire that was developed for the larger study on pesticide exposure and the home environment. These were returned by post to the Centre of Public Health Research in a prepaid, self-addressed envelope.

Upon return of these questionnaires, 180 participants from the Wellington region, and 55 from Hawke's Bay³ were asked to take part in Phase 2 of the study. Prior to beginning Phase 2 data collection, parents were sent a letter regarding their child's participation in the next part of the study. This letter described again what the

³ As Hawke's Bay data collection is still underway at the time of writing this thesis, this number will be larger in future studies released by the larger research team.

second phase of the study entailed and also reminded parents of their right to withdraw if they no longer wished to participate. Three declined continued participation. Also prior to Phase 2 beginning, two members of the research team met with staff at each school to discuss the schools involvement in the second part of the study.

Phase 2. Phase 2 data collection began in December of 2014 and continued through to August of 2015. Phase 2 involved collection of teacher information about the child's emotional and behavioural functioning and EF within the school environment. Teachers were delivered information and consent forms via the school and, on receiving their consent, were sent the BASC-2 and BRIEF teacher forms which were also returned via post to the Centre of Public Health Research in a prepaid, self-addressed envelope. Phase 2 data collection also involved gaining objective information about the child's performance, including a neuropsychological assessment. Each assessment took approximately 1 hour. These tasks were all administered by three Doctoral Students in Clinical Psychology with experience in administering norm-referenced tests.

Participants

Participants in the current study were 241 students aged 5-12 (mean age 8.31). As above, these participants were from eleven primary schools in two New Zealand (NZ) cities. Participants for the current study were originally to be recruited from one large NZ city, however as too few participants had complete data sets (parent, teacher and neuropsychological assessment data), data were collected from a second NZ city.

Schools were approached initially via email contact, and then phoned three days later so that the study could be discussed with the Principal. If the Principal was

interested in participation, a meeting was coordinated to discuss the study in more detail and adapt practical aspects of the larger study to meet the needs of the individual school. Following this meeting, the Principal then took the information through their internal procedures which typically involved taking it to the Board of Trustees for approval before commencement of the study.

Schools were approached based on decile⁴, proximity to the cities, and size of school roll. The rationale for this was to make sure that a cross-section of the general population was represented. Decile ratings in NZ represent the socioeconomic status of families in that geographic region. As it turned out however, predominantly higher decile schools agreed to participate in the study, with lower decile schools citing participation in other research projects and less experienced staff as reasons for declining. The children involved in the current study, and this aspect of the wider study, attended schools ranging from decile 3 to 10 with a rating of 3 indicating a moderate proportion of children from low socio-economic communities and a rating of 10 suggesting that the majority of children attending the school were from high socio-economic communities (Ministry of Education (MOE), 2015). Characteristics of participating schools are depicted in Table 1.

Proximity to cities was included primarily for pragmatic reasons when approaching schools, as data collection involved considerable travel to and from schools. However, as it turned out, as many of the schools closer to town were already involved in other research projects, a number of schools outside of the main city centres were approached.

⁴ Lower decile indicates lower socio-economic status

The school rolls of participating schools ranged from a roll of 112 to 596. In New Zealand a school roll of 200 or less is considered small (Collins, 2004), suggesting that the majority of participating schools in the wider study were of small to moderate size. The number of participants per school ranged from 9 to 49 with a total sample of 243 (121 male). Prior to data analyses being conducted, two cases were excluded from the data that had been assessed. Exclusion criteria stipulated that children with intellectual disabilities would not be excluded from participating in the assessments if they wished but would be excluded from data analyses. Both individuals had been diagnosed with multiple developmental disorders including autism and developmental delay and one of these children was also non-verbal. These exclusions were consistent with the NEPSY-II standardisation sample reported in the Clinical and Interpretive Manual (Korkman et al., 2007). Following these exclusions, 241 participants were included in analyses. Gender composition, while varied from school to school, was representative of the combined school population. School roll and gender composition of each school are also depicted in Table 1.

As can be seen in Table 1, there was a great deal of variance in demographic composition across schools. Unfortunately, our sample was not representative of the overall school population in terms of ethnicity, with a higher proportion of NZ European participants and a lower proportion of Maori participants than the overall population (Education Counts, 2017). There was also an uneven distribution of schools as a function of SES, with 108/241 participants coming from decile 9 or 10 schools and only 39 from low decile 3 to 5 schools. As above, a concerted effort was made to recruit lower decile schools however this was not the case. Implications are discussed in Chapter 10.

Demographic data for the school population was taken from the school's most recent report as documented on the Education Review Office website (ERO, 2015).

Demographic data for the sample was collected from parents and guardians as part of the pesticide questionnaire in order to compare gender and ethnic distributions, and is depicted in Table 1.

Table 1.

Demographic composition of sample by school.

School	Decile ^a	Roll	Sample N (male)	Ethnic composition of sample					
				NZ Euro	Maori	Other ^b	NZ Euro & Maori	NZ Euro & Other	Maori & Other
A	10	201	18 (6)	14	0	1	1	2	0
B	9	185	9 (2)	7	0	1	0	1	0
C	3	316	11 (6)	7	1	1	1	0	1
D	7	390	47 (23)	27	1	2	5	12	0
E	10	112	10 (5)	7	0	3	0	0	0
F	8	327	30 (19)	26	0	1	2	0	1
G	5	270	28 (13)	18	0	2	2	5	1
H	9	149	19 (8)	15	0	2	1	1	0
I	10	178	18 (10)	16	0	0	0	2	0
J	8	302	17 (10)	16	0	1	0	0	0
K*	10	596	34 (19)	31	0	0	2	1	0
Total		3026	241 (121)	184	2	14	14	24	3

^aDecile ratings reported are taken from ERO (2015). Decile ratings changed in 2015 and it is these that have been reported. ^bOther includes Dutch, Dutch-French, Bantu-Cameroon, West Africa, Cambodian, Chinese, Pacific, French, Indian, Fijian Indian, Greek Cypriot, Irish, Sri Lankan, European, Asian Indian, Bhatanese (Nepali), Japanese, Serbian, American, South American (Chile), Samoan, German, Palestinian and British. Ethnicity was indicated by parents on the demographic questionnaire. In NZ, individuals can indicate multiple ethnicities, thus people can indicate multiple ethnicities under 'Other' and more ethnicities may be presented in this column than participants.

*School K was located in Hawke's Bay. All other schools were located in Wellington.

Measures

Behaviour Assessment System for Children, Second Edition (BASC-2). The BASC-2 (Reynolds & Kamphaus, 2004) is a system of brief rating scales which assess emotional and behavioural functioning. This system consists of a Teacher Rating Scales (TRS), a Parent Rating Scale (PRS), a Self-Report of Personality, a Student Observation System, and a Structured Developmental History (SDH) form.

The BASC-2 was initially selected for this study over similar measures such as the Child Behaviour Checklist (CBCL; Achenbach & Rescorla, 2001) and the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) as it is more recent, has strong reliability and validity estimates, is used in NZ already, and also due its unique addition of the SDH. Ultimately however, the decision was made not to use the SDH as the larger study required more targeted questions regarding pesticide exposure as well as developmental history. For the purposes of both the current and the larger study, the TRS and PRS were selected to provide a general overview of the child's emotional and behavioural functioning within the home and school contexts.

Due to slight differences in NZ and American English, both questionnaires were reviewed for cultural appropriateness by the research team. Some items contained a word that was less common in NZ than America. One example of this was the use of the word *grade*, which is less common in NZ English. To maintain the integrity of the form, a sticker was attached to the front cover clarifying items that may be confusing (see appendix B).

The BASC-2 TRS contains 139 items for which there are 4 response options: never, sometimes, often, and almost always. Items load onto domains of externalising problems (aggression, conduct problems, and hyperactivity), internalising problems (depression, anxiety and somatization) adaptive skills (adaptability, social skills and leadership) and school problems (study skills, attention problems, learning problems). The BASC-2 PRS is completed in the same manner, but differs in that it contains 160 questions regarding externalising and internalising problems, adaptive skills, and activities of daily living. Each questionnaire takes approximately 10-20 minutes to complete (Reynolds & Kamphaus, 2004).

Once completed and returned, the PRS and TRS forms were mostly hand scored by the author in accordance with the scoring procedures reported in the manual (Reynolds & Kamphaus, 2004). In October 2015 the research team acquired the scoring software for the BASC-2, and final scoring was completed using this software. Approximately 20 forms that had been hand-scored were entered into the computing software to ensure that scoring was consistent across formats. All scoring was identical across formats.

Scoring on the BASC-2 involves converting the raw scores for each domain and composite score to *T*-scores that are adjusted for age and gender. In this way, the child's *T*-scores on a particular domain should ultimately reflect their behaviour compared to other children of the same age and gender.

Critical Items were reviewed to determine if there were any items indicated by parents that may be of clinical concern. Any critical items or elevated scores that were identified were discussed with a Registered Clinical Psychologist, and any suspected risk was followed up with the parent/guardian of the child. If parents identified that they were concerned about the items they were offered both contact numbers of free support services and to be put in contact with the Psychology clinic on campus. Parents were also offered either the number for or to be put in touch with the Massey University Psychology Clinic, a private Psychology Clinic run through the Wellington Campus, though they were informed that this service would have fees associated with it. The prevalence of these items was higher than expected, therefore the decision was made to include an evaluation of Critical Items in the analysis.

For forms that had been computer-scored, content scales could also be identified for analysis. Two of the content scales available on the BASC-2 ASSIST Plus software are Emotional Self-Control and Executive Functioning. These content scales were

included for analysis to determine the extent to which these scales related to performance on EF and Social Perception tasks, but also the extent to which they corresponded with domains on the BRIEF. The BRIEF and performance tasks will be described shortly.

Missing data were accounted for in the scoring procedures documented in the BASC-2 manual (Reynolds & Kamphaus, 2004). The BASC-2 allows for no more than two items per scale that are unable to be scored (Reynolds & Kamphaus, 2004). More than two items would suggest that the scale was no longer valid. Invalid scales were not included in analyses.

The BASC-2 is reported to have sound psychometric properties, with internal consistency estimates ranging from mid .80s – mid .90s for composite scales (e.g. externalising problems) and .80 -.89 for individual scales (e.g. aggression) (Tan, 2007). It is also reported to have high criterion-rated validity (Tan, 2007).

Behaviour Rating Inventory of Executive Function (BRIEF). The BRIEF (Gioia et al., 2000a) was selected as it is the only subjective rating scale used in NZ that provides an overall index of EF within the home and school environments. It is appropriate for children aged 5 – 18 years. Like the BASC-2, it is a rating scale with both parent and teacher forms and takes approximately 10-20 minutes to complete. Both forms consist of 86 questions relating to 8 domains of EF: inhibition, the ability to shift cognitive sets, emotional control, initiation, working memory, the ability to plan/organise, organisation of materials and monitoring ability. Unlike the BASC-2 however, the BRIEF only uses three response options: never, sometimes, and often.

The BRIEF forms were also reviewed for cultural appropriateness by the research team. As with the BASC-2 forms, a sticker was attached to the front cover to clarify potentially confusing items (see appendix B).

Once completed and returned, the parent and teacher forms were primarily hand scored following the scoring procedures reported in the manual (Gioia et al., 2000a), though as with the BASC-2, computing software became available for use from October 2015 meaning that remaining forms could be computer-scored. As with the BASC-2, missing data are accommodated for in the scoring procedures. Similarly the BRIEF allows for no more than two items per scale that are unable to be scored, and no more than 14 items overall (Gioia et al., 2000a). Invalid scales were not included in analyses.

Scoring of the BRIEF also involves converting raw scores to *T*-scores in order to determine how the child's behaviour compares to age- and gender-matched peers. The BRIEF is reported to have strong internal consistency ranging from .80-.98, and test re-test reliability averaging from .81 (parent) to .87 (teacher) (Roth, Isquith & Gioia, 2014). Few measures similar to the BRIEF are available but more recent research has found concurrent validity between the BRIEF and the attention problem scale of the BASC-2, and with various ADHD diagnostic scales (Roth et al., 2014).

Neuropsychological Assessment Measures

Tests were selected by three clinical psychology students and a Registered Clinical Psychologist. Different neuropsychological assessments were reviewed that were appropriate to use with children aged 5 – 12 and were similar to those used in previous studies looking at pesticide effects on cognitive function. Subtests for the larger study were selected across domains to provide an overview of cognitive status, and are depicted in Table 2. Following this review, thirteen subtests were selected from

the NEPSY-II, WISC-IV, and TEA-Ch assessment systems. As the primary focus of the current research was performance on the EF and Social Perception domains, the subtests selected from the WISC-IV and TEA-Ch assessment systems for the larger study were not relevant to the current study and therefore will not be described here. Further information on these systems will be reported elsewhere.

Final test selection took into account the closeness of the test to the cognitive domain it was designed to assess, the age range it was intended for and the length of time taken to administer the task. As testing took place at the child's school, every effort was made to complete as comprehensive an assessment as possible in a limited time to minimise disruption to the child's schooling. The entire assessment took approximately one hour to administer.

The order in which subtests were administered was also reviewed due to concerns about practice and interference effects. A 20 minute delay was placed between the Memory for Faces subtest and the Memory for Faces Delayed subtest to reduce the effects of practice on this task, and tests such as AR and ToM which involved visual presentation of children's faces were administered after both Memory for Faces tests to eliminate possible interference confounds.

Table 2

Subtests administered during neuropsychological assessment

Cognitive Domain	Subtest Administered	Approximate administration time (mins)
Social Perception	AR - NEPSY-II	6
	ToM - NEPSY-II	8
Attention and EF	Animal Sorting - NEPSY-II	6
	Inhibition - NEPSY-II	7
	Auditory Attention and Response Set - NEPSY-II	8
	Sky Search -TEA-Ch	4
Working Memory	Digit Span forward and backwards - WISC-IV	3
Memory	Memory for Faces - NEPSY-II	4
	Memory for Faces Delayed - NEPSY-II	3
	Narrative Memory - NEPSY-II	4
Sensorimotor	Finger Tapping - NEPSY-II	1
Processing Speed	Coding - WISC-IV	2
Language	Comprehension of Instructions - NEPSY-II	4

Developmental NEUROPSYCHOLOGICAL ASSESSMENT (NEPSY-II). The NEPSY-II (Korkman et al., 2007) provides an objective measure of the child's cognitive status. It consists of 36 subtests across 6 domains: Attention and EF, Language, Memory and Learning, Sensorimotor, Social Perception and Visuospatial Processing. NEPSY-II administration allows for subtests to be selected according to the purpose of the examination, without having to administer the entire assessment system (Korkman et al., 2007).

The NEPSY-II is reported to have robust inter-rater reliability ranging from 93-99%, adequate to high internal consistency, good discriminative validity and concurrent validity with the WISC-IV (Davis & Matthews, 2010).

As previously described, the key component of the assessment for the current study was the child's performance on the Social Perception and EF domains. The Social Perception domain consists of two subtests, ToM and AR, which were selected along with two EF subtests, Inhibition and Animal Sorting.

Social perception subtests

ToM. The ToM subtest comprises of a range of questions designed to assess a child's understanding of another's thoughts, beliefs and emotions, their understanding of intention, deception and pretence, and their ability to apply social context to determine appropriate emotion or affect (Korkman et al., 2007). Questions or scenarios are posed through vignettes and most are accompanied by pictures. A contextual task is also used to determine the participants understanding of how the child in the picture would feel in a given social context. For instance, one picture illustrates a child who has fallen off her bike, and the child is asked to choose from four pictures depicting various emotions the one that most closely resembles how the child in the picture is feeling.

This subtest was reviewed for cultural appropriateness due to the use of American names and activities that would likely have interfered with children's understanding of the questions being asked. Again, questions were reviewed by the research team which consisted of three clinical psychology students in their 1st and 2nd years of training and a Registered Clinical Psychologist. Questions were also reviewed by another Clinical Psychologist with no links to the study. This practitioner also had children within the appropriate age range for the test. Revisions can be found in Appendix B.

AR. In this task, participants are assessed on their ability to recognise affect in other children. Participants are visually presented with sets of photographs displaying happy, sad, angry, disgusted, fearful or neutral faces and across four different tasks are asked to select photographs where the children look as if they are displaying the same emotion (Korkman et al, 2007).

EF subtests

Inhibition. This subtest consists of three different tasks, all of which are timed to assess the fluency with which the child can complete the task. The first, a naming task, requires the child to look at a series of shapes (black and white circles and squares) and name them as fast as possible. The second, an inhibition task, requires the child to inhibit the automatic response to say the shape's correct name and instead say the opposite shape's name. The third, a switching task, requires the child to inhibit the shape's correct name when the shape is white and report the shape's correct name when the shape is black. These three tasks are all then repeated using an alternate form. The alternate form consists of a series of arrows, where the child reports the direction the arrow is facing (up or down) and the final score for each task is a combined scaled score of performance on both the shape and arrow conditions.

Animal Sorting. This subtest was included for children aged seven and up, and involves the child sorting cards into two groups of four cards using their own sorting criteria (Korkman et al., 2007). There are 13 possible correct sorts which the child could use to make groups. This subtest was included as it tests a number of EFs, including initiation, switching, and problem-solving ability.

NEPSY-II Scoring

Scoring for the NEPSY-II was completed following the scoring procedures in the manual (Korkman et al., 2007). Children's raw scores on the NEPSY-II tasks were converted to scaled scores that are adjusted for the child's age at testing and the task being measured. The NEPSY-II Scoring Assistant and Assessment Planner was used to enter the data and determine scaled scores for each subtest.

Changes to Test Administration

Originally, 5 and 12 year olds were not included in the study as they were outside the age range of norms for the BASC-2, however on discussion with the school Principals in the beginning stages of data collection this decision was revised as Principals felt that if the school was to be part of the study they would like to offer everyone the opportunity to participate. As the majority of those 5 and 12 year olds who participated in the study were within months outside of the normative age range it was felt that they were close enough in age to be included in the study. Also given differences in school entry across cultures, with NZ children entering formal schooling on their 5th birthday and American children beginning formal schooling at the start of the year, NZ children are likely to have had more exposure to formal schooling at the time the tests were administered. Therefore, it was decided that it would make more sense to administer the child version of the BASC-2 as opposed to the preschool version despite differences in age.

On scoring the BASC-2 PRS and TRS forms, children age 5 were scored using the norms for a 6 year old and children aged 12 were scored using the norms for a child aged 11. Independent sample *t*-tests with the BASC-2 PRS and TRS showed no significant difference when comparing children who were outside of the target age

range, and thus the full sample was included in analysis. BRIEF norms extend from 5-13 and the selected subtests from the NEPSY, with the exception of Animal Sorting (age 7-16) and Inhibition (age 5-16) have norms available from age 3 so age was not a factor in analyses of these measures. Children aged 5 and 6 were not administered the Animal Sorting task.

Inter-Rater Reliability

Prior to data collection, the researchers familiarised themselves with the materials and manuals, and practiced administering them together, ensuring agreement over scoring. Then for the first six participants in the study, the experimenters shadow marked each other and inter-rater reliability was calculated by adding up the number of times the experimenters agreed, and dividing this by the total number of times they agreed and disagreed. On average an inter-rater reliability score of 96.84% was found, suggesting there was consistency across examiners. This is consistent with NEPSY-II inter-rater reliability criteria that required examiners in their standardisation process to have an inter-rater reliability of at least 90%.

Exclusion Criteria

Exclusion criteria for the current study included that the participant understood English and was within the ages of 5 to 12. Children with significant intellectual disabilities were evaluated on a case by case basis for the purposes of analysis but were not excluded from participating in the assessments if they wished. As we did not have

complete parent, teacher and neuropsychological assessment data from each participant, sample sizes for analyses in each study differ as a function of what data was available.

Testing Environment

Children were assessed individually in a quiet space at their school and short breaks were provided when required. Assessments took place during school hours and were coordinated with the school to be as minimally disruptive as possible. Testing sessions were arranged with the contact person from the school, usually the school Principal, to ensure that testing times would be suitable for the school and that the child and their teacher were given notice before we arrived. In general, school visits were arranged so that participating children could be assessed one after the other. Due to space restrictions, testing environments were not consistent across schools. Testing environments included resource rooms, libraries, and the staffroom. If a choice of spaces was available, the room that was least likely to be disrupted was selected.

The physical setup of the room also differed marginally but adhered as much as possible to the physical setup outlined in the NEPSY-II manual (Korkman et al, 2007). A table or desk was set up in the middle of the room with a chair on either side. The child sat on one side and the researcher on the other. The stimulus book was placed in front of the child. Two pencils and a stopwatch were placed on the table. The laptop computer, animal sorting cards and ToM boxes as well as the stationery items for when the child had completed the assessments were placed beside the researcher but where they would not distract the child during other tasks.

At the beginning of each neuropsychological evaluation, rapport was developed with the child and the basic purpose of the session was relayed in language appropriate

for their developmental age. Assent was then gained from the child before proceeding with the tasks.

Compensation

As compensation for their efforts in supporting the research projects, each school was offered \$200 to use in a way that would be most beneficial for them. This was discussed with the Principal and was offered either in the form of a prize draw for individual participants or alternatively to be used toward school resources, such as sports equipment or book vouchers. Individual students were invited to choose a stationery item as a reward for their efforts following the neuropsychological assessments. Compensation was also provided in the form of a science activity at the school which was designed to teach children the basics about the brain and how they can protect their brains; for example, wearing a helmet when riding a bike, and eating healthily. This science lesson involved many interactive activities and viewing of a model brain.

Confidentiality

Results for individual children were kept confidential. Questionnaires were stored in a locked filing cabinet in CPHR and all results were filed in password protected databases that were only accessible to researchers associated with the current and larger studies.

Feedback

As part of the larger study a summary of the participants' scores was made available to parents upon request. Parents were informed at the time of request and on

the feedback form that these tests were selected for a specific purpose and that it was not a comprehensive neuropsychological assessment. They were also informed that due to the disruptive nature of testing in schools and other variables that we could not control for, the report may not accurately reflect their child's typical performance. Any summaries made available to parents were completed by a registered clinical psychologist with extensive experience in neuropsychological assessment. They were not completed by the author.

CHAPTER SIX

Preliminary Analyses

Overview

The results of this thesis are split into four sections; preliminary analyses and Studies One, Two, and Three. Preliminary analyses relate to all three studies, and are reported first. For ease of readability, each study will begin with a brief overview of the research question being assessed, followed by the results, and a discussion of the findings. A summary of all findings will then be discussed in the conclusion.

First the scoring and collation of data will be discussed, followed by the results of analyses to confirm that data met the assumptions for use of parametric statistics in analyses. Internal consistency of the measures are described to determine if the estimates provided by the test manuals are still valid with our sample. Finally, examiner effects were evaluated to determine whether results differed on the objective assessment tasks as a function of examiner. Data analyses were conducted using IBM SPSS Statistics version 22 for Windows.

Data Collation

All data were collated and scored following standardised procedures in the test manuals. Composite scores were used for analysis. Scores were then aggregated across participants to gain a mean score for each domain. Data were grouped for analysis to protect anonymity of individual participants. Due to the large number of participants involved in the large study and the difficulties of gathering parent, teacher and neuropsychological assessment data from each participant, at the time of data analyses the sample sizes indicated for each analysis differ as a function of the data that was available. The *n* available for each comparison is shown in Table 3.

Table 3.

Available n for comparison across measures

	PRS	TRS	PF	TF	NEPSY	AS
PRS	216					
TRS	99	106				
PF	213	103	222			
TF	99	104	104	106		
NEPSY-II	97	94	101	93	118	
AS	76	74	77	73	89	89

Note. PRS = BASC-2 Parent Rating Scale; TRS = BASC-2 Teacher Rating Scale; PF = BRIEF Parent Form; TF = BRIEF Teacher Form; NEPSY-II = ToM, AR and Inhibition subtests; AS = Animal Sorting subtest from NEPSY-II. *N* is low for AS as it was not administered to children under 7 years of age.

As with any large study with the sheer quantity of data being analysed, there were concerns regarding human error in scoring, entering and analysing data. For the current research a large portion of questionnaire data were hand-scored and entered by the author as computer-scoring was not available at the time. These were then double-checked by the author, and cross-checked by CPHR staff who also entered data into a separate database. Any errors were referred back to the author to check. In this way the author was able to ensure that data were scored and entered consistently but also ensure that there were procedures in place to minimise possibility of errors.

Testing Assumptions of Normality

The first step in conducting preliminary analyses was to determine whether the data met assumptions for use of parametric statistics. Across individual domains on the BASC-2 and BRIEF measures, skewness ranged from -.85 to 2.75 with kurtosis ranging from -.80 to 10.54, indicating that individual domains were not normally distributed (Tabachnick & Fidell, 2013). Parent-rated forms (skewness: -.34 – 1.87; kurtosis: -.67 –

4.64) were less skewed than teacher-rated forms (skewness: $-.35 - 2.75$; kurtosis: $-.79 - 10.54$), potentially due to sample size, and composite scores were less skewed than individual domains. Objective performance on the NEPSY-II tasks was normally distributed with skew ranging from $-.85$ to $.81$, and kurtosis ranging from $-.80$ to 1.16 .

Given that the subjective measures do not appear to be normally distributed, non-parametric statistics were performed alongside parametric statistics (Tabachnick & Fidell, 2013), though in general, findings did not differ between parametric and non-parametric statistics. As parametric statistics are more robust however (Tabachnick & Fidell, 2013), it is these that will be reported throughout the results. All non-parametric analyses are included in Appendix C.

Internal Consistency

Previous studies utilising the BASC-2 have reported internal consistency estimates ranging from mid .80s – mid .90s (Tan, 2007). In the current study the Cronbach alpha coefficients for problem and adaptive behaviours on the PRS was .89 and .9 respectively. On the TRS, the Cronbach alpha coefficients for problem and adaptive behaviours were .90 and .92 respectively.

Internal consistency on the BRIEF is reported to range from .80-.98 (Roth et al., 2014). In the current study the Cronbach alpha coefficient on the parent form was .94 and .96 on the teacher form.

Examiner Effects

As well as ensuring inter-rater reliability prior to beginning data collection, data were checked prior to analyses to ensure no examiner effects could be accounting for variance in NEPSY-II performance. Mean scores for each NEPSY-II subtest by

examiner are depicted in Table 4. No significant differences were found between examiners on any of the subtests except ToM. On the ToM subtest, there were significant differences between examiners 1 and 3 ($t = 2.433$; $p = .017$; 95% CI [.233, 2.308]) and examiners 2 and 3 ($t = .3702$; $p < .001$; 95% CI [.900, 3.007]). There were no significant differences between examiners 1 and 2. It is worth noting that the n for this task differed across examiners (Examiner 1: $n = 40$; Examiner 2: $n = 24$; Examiner 3: $n = 54$). It is possible therefore, that Examiner 3 saw more variety in ToM abilities than the other two examiners. It is also possible, with this being a relatively small sample, that this difference would not have been significant with a larger n . Mean scores for each subtest by examiner are depicted in Table 4.

Table 4.

Mean scores on NEPSY-II subtests by examiner

	Examiner 1	Examiner 2	Examiner 3
	Mean	Mean	Mean
	(s.d.)	(s.d.)	(s.d.)
AR	9.78	9.88	9.83
	(2.32)	(2.58)	(3.00)
ToM	10.40	11.08	9.13
	(2.26)	(1.79)	(2.80)
Animal Sorting	8.32	8.89	7.83
	(3.57)	(3.32)	(2.80)
Naming	10.65	11.50	11.13
	(3.22)	(3.53)	(3.01)
Inhibition	10.30	10.79	10.61
	(3.07)	(2.84)	(3.40)
Switching	11.84	10.68	11.08
	(2.34)	(3.15)	(2.52)

$N = 118$

CHAPTER SEVEN

Study One

Overview

The purpose of this study is to evaluate the cultural appropriateness of the BASC-2, the BRIEF, and the NEPSY-II subtests for a NZ sample, and is reported in three parts. The first section of this study provides a preliminary evaluation of the ‘Critical Items’ listed in the BASC-2, and the prevalence with which these items occur in a NZ sample. Part Two compares scores on each measure to the US standardisation samples using one-sample *t*-tests. Finally, Part Three reports the results of comparisons across gender and decile. The study concludes with a discussion of the results for each section and overall findings.

Part One

Overview. The ‘Critical Items’ identified in the BASC-2 refer to specific problem behaviours that may be clinically important but may otherwise be obscured in profile interpretation (Reynolds & Kamphaus, 2004). Reynolds and Kamphaus report that a response of “sometimes”, “often”, or “almost always” on any of these items may warrant follow-up. Many of these items are also clinically important to follow up as they may indicate risk of the child to themselves or others, neurological concerns, or other mental health concerns that may not otherwise be evident from domain or composite scores. Individual items are not reliable predictors of behaviour (Reynolds & Kamphaus, 2004), and as such, critical items have not been reported in previous research. However, due to likely differences in the expression of emotional and behavioural problems in NZ compared to the USA, determining the prevalence of these problem behaviours in NZ may be of benefit to people administering this tool. In other

words, in order to know what weight to assign these individual items within a NZ cultural context, it is important to know how frequently they occur within the population. Though there is no earlier research to support the hypothesis it is expected that prevalence will be low given that this is a non-clinical sample. It is also expected that prevalence on some items will differ as a function of the child's age.

Results. The prevalence of individual Critical Items on the BASC-2 PRS and the BASC-2 TRS are shown on Table 5. Interestingly, 155 (72.1%) parents indicated that either their child sometimes, often or almost always ate too much (47.0%) or too little (25.1%) and 158 (73.5%) of parents endorsed that their child is easily annoyed by others sometimes, often or almost always. Items such as 'Sleeps with parents', and 'Wets bed' were reasonably common but significantly correlated with age (Sleeps with parents $r = -.20$; $p = .003$; Wets bed $r = -.17$; $p = .013$), suggesting that, as would be expected, these behaviours declined with age. In contrast, other items increased in likelihood with age (Is easily annoyed by others $r = .20$; $p = .003$; Says "I want to die" or "I wish I were dead" $r = .18$; $p = .007$; and Says "I want to kill myself" $r = .20$; $p = .004$)⁵. Remaining items were not significantly correlated with age.

⁵ Follow up phone calls were made to any parent who endorsed either that their child says "I want to die" or "I wish I were dead", or if they endorsed that their child says "I want to kill myself". The general content of these calls is discussed on pages 62 and 63.

Table 5.
Frequency of PRS and TRS Critical Items

Item	PRS				TRS			
	Never n (%)	Sometimes n (%)	Often n (%)	Almost Always n (%)	Never n (%)	Sometimes n (%)	Often n (%)	Almost Always n (%)
Is easily annoyed by others	56 (26.2)	139 (65.0)	18 (8.4)	1 (.5)	60 (56.1)	41 (38.3)	6 (5.6)	1 (1)
Eats too much	113 (53.3)	82 (38.3)	17 (7.9)	1 (.5)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Falls down	155 (72.4)	58 (27.1)	1 (.5)	0 (0)	95 (89.6)	11 (10.4)	0 (0)	1 (1)
Sleeps with parents	154 (72.0)	52 (24.3)	6 (2.8)	2 (.9)	98 (94.3)	5 (4.8)	0 (0)	1 (1)
Eats too little	159 (74.6)	43 (20.2)	11 (5.2)	0 (0)	105 (99.1)	1 (.9)	0 (0)	0 (0)
Has toileting accidents	170 (79.4)	41 (19.2)	3 (1.4)	0 (0)	94 (88.7)	10 (9.4)	2 (1.9)	0 (0)
Hits other children	177 (82.7)	35 (16.4)	2 (.9)	0 (0)	91 (87.5)	13 (12.5)	0 (0)	0 (0)
Wets bed	170 (79.4)	31 (14.5)	8 (3.7)	5 (2.3)	102 (95.3)	4 (3.7)	1 (.9)	0 (0)
Bullies others	188 (87.9)	26 (12.1)	0 (0)	0 (0)	105 (99.1)	1 (.9)	0 (0)	0 (0)
Threatens to hurt others	189 (88.7)	24 (11.3)	0 (0)	0 (0)	94 (88.7)	6 (5.7)	3 (2.8)	3 (2.8)
Says "I want to die" or "I wish I were dead"	190 (88.8)	22 (10.3)	2 (.9)	0 (0)	103 (97.2)	2 (1.9)	0 (0)	1 (.9)
Has eye problems	180 (84.1)	23 (10.7)	4 (1.9)	7 (3.3)	106 (100)	0 (0)	0 (0)	0 (0)
Has a hearing problem	195 (91.5)	16 (7.5)	1 (.5)	1 (.5)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Throws up after eating	202 (94.4)	12 (5.6)	0 (0)	0 (0)	106 (100)	0 (0)	0 (0)	0 (0)
Says "I want to kill myself"	205 (95.8)	9 (4.2)	0 (0)	0 (0)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Is cruel to animals	206 (96.3)	8 (3.7)	0 (0)	0 (0)	106 (100)	0 (0)	0 (0)	0 (0)
Eats things that are not food	207 (96.7)	5 (2.3)	2 (.9)	0 (0)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Sees things that are not there	208 (97.2)	5 (2.3)	1 (.5)	0 (0)	102 (95.3)	5 (4.7)	0 (0)	0 (0)
Runs away from home	210 (98.1)	4 (1.9)	0 (0)	0 (0)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Sets fires	211 (98.6)	3 (1.4)	0 (0)	0 (0)	106 (100)	0 (0)	0 (0)	0 (0)
Hears sounds that are not there	211 (98.6)	2 (.9)	1 (.5)	0 (0)	106 (99.1)	1 (.9)	0 (0)	0 (0)
Has seizures	212 (99.1)	1 (.5)	1 (.5)	0 (0)	106 (100)	0 (0)	0 (0)	0 (0)

N = 214 due to missing items that were unable to be collected.

In our study, 24 (11.2%) individuals were reported to sometimes or often say “I want to die” or “I wish I were dead” and nine (4.2%) were reported to sometimes say “I want to kill myself”. Eight of the individuals who were reported to have said “I want to kill myself” also were reported to have said “I want to die” or “I wish I were dead”.

Returned forms with endorsed Critical Items were discussed with the author’s supervisor who is a Registered Clinical Psychologist. It was decided that she would phone the first five parents to enquire about the context in which these behaviours were occurring, whether the parents were concerned about the behaviour, and whether they required additional support. Being reassured that there were no pressing concerns during these phone calls, a protocol was drafted for making calls and future follow up calls were made by the author, a trainee Clinical Psychologist, and discussed in supervision.

Phone calls were made to the parents of all 25 individuals identified who endorsed one or both of these items. An additional 8 parents were also contacted who had endorsed Critical Items such as ‘Hears sounds that aren’t there’, ‘Is cruel to animals’ and ‘Runs away from home’, making a total of 33 follow up calls. Of those, 7 could not be contacted despite multiple attempts. Of the 26 parents who were contacted, 7 indicated concern, though 5 of these were already receiving support from Mental Health Services. Two parents who had concerns were provided with contact details for Youthline and Lifeline (two free 24/7 contact numbers that parents can call if they have concerns) and one parent also accepted the phone number for the Massey University Psychology Clinic. Nineteen parents were not concerned about the behaviour, stating they understood these statements to be something their child said to get a reaction and it was not concerning for them. Implications of these findings are addressed below.

Discussion. As far as this author is aware, this is the first study to identify prevalence of the BASC-2 Critical Items. Given that a non-clinical sample was used, it was hypothesised that the prevalence of these items would be low as they are thought to be indicators of serious concern and not likely to be highly prevalent within the general population. With no previous studies to compare to however, it was difficult to predict what the spread would look like. The results of this analysis revealed that the prevalence was higher than expected on a number of items. The number of parents who endorsed death-related Critical Items, for instance, was higher than expected, particularly when considering Ministry of Health (MOH, 2012) estimates that the prevalence of emotional and behavioural disorders for NZ children is approximately 4%.

The results indicated that again, it was important to follow up these items to gather the context of the behaviour. Though the majority of the time the behaviour was not a concern, over a quarter of parents who endorsed that their child either said “I want to die”, “I wish I were dead” or “I want to kill myself” (26.9%) did report concern. While in this study, most of those had already sought the assistance of Mental Health Services, the two that had not reported that they had “not really thought about it” until being asked to complete the form. Thus in future research, being aware of Critical Items and following up with parents may be an important ethical concern to consider.

As expected, age was a factor in the prevalence of some items, such as wetting the bed. While this behaviour may not be concerning for a child of six, it is more concerning if the child is 11. Therefore, it is important to consider context and not just the item itself. Age was taken into account when evaluating whether to follow up items with the parent, and context of items was discussed with the parent if a follow up call was made. Critical Items on the TRS were not followed up, due in part to the fact that those identified had already been followed up with the parent but also due to the fact

that this had not been discussed in the ethics application or with parents when they consented to participate.

Part Two

Overview. Part Two compared the results from our sample to the U.S norms. As scores on the BASC-2 and BRIEF had been converted to *T*-scores as part of the scoring procedures, a mean *T*-score of 50 ($SD = 10$) reflects the mean of the standardisation sample. One sample *t*-tests were therefore selected to compare responses from our sample to a known test value mean of 50.

One-sample *t*-tests were also selected to evaluate whether the NEPSY-II Social Perception and EF measures differed from the standardisation sample. As the NEPSY-II produces scaled scores, responses in our sample were compared to test value of 10 ($SD = 3$). Based on previous research in NZ (e.g., Fernando et al., 2003; Ogden & McFarlane-Nathan, 2007) and elsewhere (e.g., Rosenqvist et al., 2017), it was expected that there would be differences in ratings and performance across cultures.

Results

Evaluation of the BASC-2. Comparisons are shown in Table 6 and depicted in Figure 3. On the BASC-2 PRS all composite scores of problem behaviours (Externalising Problems; Internalising Problems; Behavioural Symptoms Index) were significantly lower than the test mean of 50 ($p < .001$) with effect sizes ranging from $d = .37$ -.42 and the Adaptive Skills Composite was significantly higher than the test mean of 50 ($p < .001$; $d = .28$), indicating that overall our sample endorsed more positive behaviours and fewer negative behaviours than the standardisation sample. Similarly across individual domains, all problem behaviours were significantly lower than the

standard mean of 50 ($p < .001$, $d = .20 - .44$) and adaptive skills were significantly higher ($p < .001$, $d = .29 - .46$) with the exceptions of the Social Skills ($d = .00$) and Activities of Daily Living ($d = .10$) domains, for which the difference was not significant.

On the BASC-2 PRS, an additional analysis was performed on the data to determine how many participants scored within the clinically significant range across the domains. 27 individuals (11.20%) scored in the clinically significant range in one or more domains.

On the BASC-2 TRS, composite and individual domain scores of externalising problem behaviours, behavioural symptoms and school problems were all significantly lower than the test mean of 50 ($p < .001$; $d = .32 - .77$), however the Internalising Problems Composite and the Depression, Anxiety, and Somatisation domains within it were not significantly different from the standardisation means ($d = .02 - .15$). The TRS Adaptive Skill Composite and all adaptive domains were significantly higher than that of the standardisation sample ($p < .001$; $d = .47 - .74$).

Table 6

One-sample t-tests comparing BASC-2 scores from NZ sample with US test mean of 50 (SD = 10).

	BASC-2 PRS				BASC-2 TRS				
	Mean	SD	<i>t</i>	95% CI	Mean	SD	<i>t</i>	95% CI	
Hyperactivity	46.82	8.61	-5.431**	-4.33	45.84	6.652	-6.439**	-5.44	-2.88
Aggression	48.25	7.07	-3.648**	-2.70	45.96	5.970	-6.964**	-5.19	-2.89
Conduct Problems	46.11	7.65	-7.497**	-4.92	44.08	4.408	-13.816**	-6.76	-5.07
Externalising Problems	46.68	7.54	-6.468**	-4.33	45.02	5.514	-9.301**	-6.04	-3.92
Anxiety	47.10	10.13	-4.215**	-4.25	50.25	12.102	.217	-2.08	2.59
Depression	47.50	7.63	-4.824**	-3.53	48.71	9.830	-1.354	-3.19	.60
Somatisation	45.84	8.93	-6.842**	-5.36	48.55	8.917	-1.663	-3.17	.28
Internalising Problems	46.04	8.91	-6.534**	-5.16	49.07	10.555	-.906	-2.98	1.11
Atypicality	47.28	8.21	-4.865**	-3.82	46.98	6.979	-4.454**	-4.36	-1.67
Withdrawal	47.51	10.34	-3.544**	-3.87	46.99	8.460	-3.662**	-4.64	-1.38
Attention Problems	47.74	8.63	-3.856**	-3.41	44.02	7.379	-8.345**	-7.40	-4.56
BSI	46.80	7.29	-6.447**	-4.18	45.50	7.161	-6.470**	-5.88	-3.12
Adaptability	52.89	8.56	4.970**	1.74	54.22	7.999	5.428**	2.68	5.76
Social Skills	50.02	8.80	.031	-1.16	56.82	8.502	8.260**	5.18	8.46
Leadership	54.19	7.94	7.781**	3.13	56.97	8.702	8.249**	5.30	8.65
ADL's	50.92	8.24	1.639	-.19					
Functional Communication	52.71	8.45	4.711**	1.58	54.01	7.015	5.885**	2.66	5.36
Adaptive Skills	52.50	7.68	4.795**	1.48	56.04	7.377	8.427**	4.62	7.46

Note: df = 215 for PRS; df = 105 for TRS; **Significant at $p < .001$ level

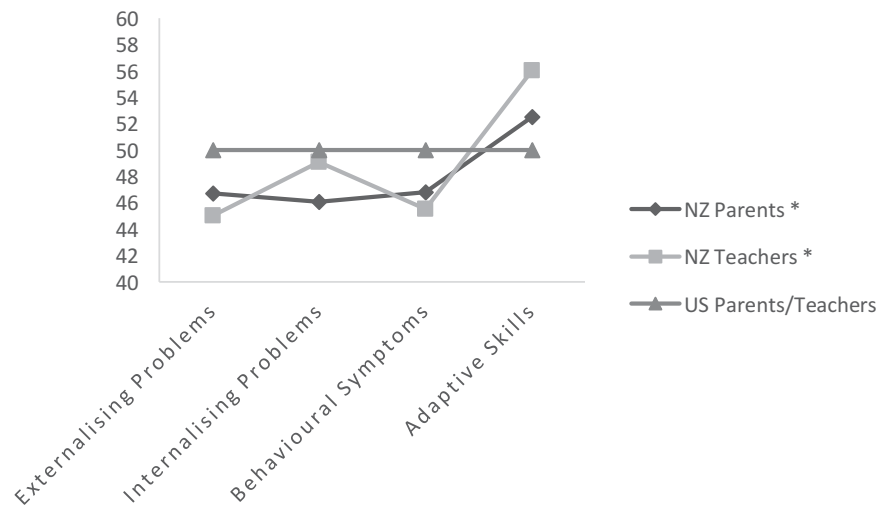


Figure 3. BASC-2 means for the total sample compared to U.S norms.

*All comparisons significant at $p < .01$ level, except Internalising Problems as rated by teachers.

Evaluation of the BRIEF. Comparisons are shown in Table 7 and depicted in Figure 4. On the BRIEF PF, the Behavioural Regulation Index was significantly lower than the standardisation mean ($p < .001$; $d = .22$), as were the Metacognition Index ($p = .009$; $d = .17$) and the Global Executive Composite ($p = .007$; $d = .17$), though the effect sizes were small. The Inhibit, Shift, and Monitor domains were significantly lower than the standardisation mean (Inhibit: $p < .001$, $d = .23$; Shift: $p = .004$, $d = .19$; Monitor: $p < .001$, $d = .37$), as was the Initiate domain ($p = .025$; $d = .14$), however across the remaining domains (Emotional control; Working Memory; Plan/Organise and Organisation of Materials) there were no statistically significant differences between the response of parents in our sample compared to those of the standardisation sample.

On the BRIEF TF there were significant differences on the Inhibit ($p = .016$; $d = .19$) and Working Memory domains ($p = .025$; $d = .20$) when compared to the

standardisation mean, though the effect sizes were negligible. Across all other domains and composite scores there were no significant differences.

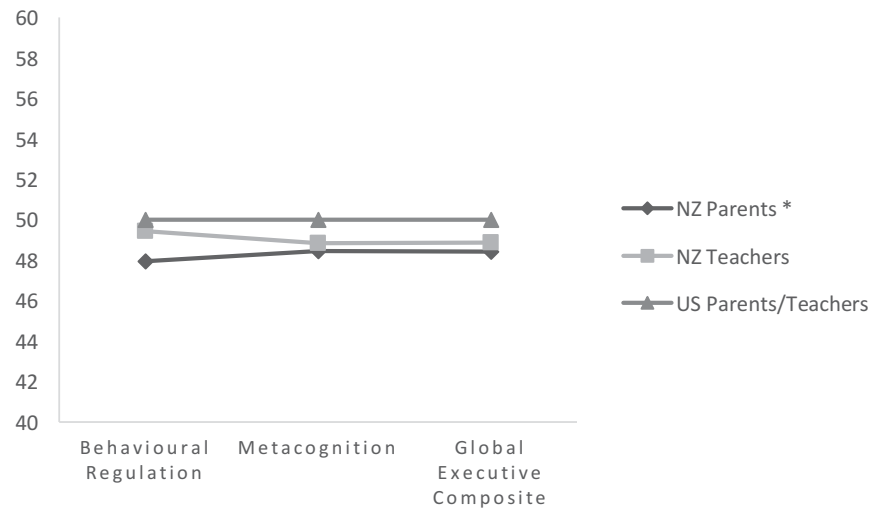


Figure 4. BRIEF means for the total sample compared to U.S norms.

*All comparisons significant at $p < .01$ level

Table 7
One-sample t-test comparing BRIEF scores from NZ sample with US test mean of 50 (SD = 10).

	BRIEF PF				BRIEF TF				
	Mean	SD	<i>t</i>	95% CI	Mean	SD	<i>t</i>	95% CI	
Inhibit	47.96	7.949	-3.833**	-3.08	-0.99	6.861	-2.452*	-2.94	-0.31
Shift	48.08	9.830	-2.923**	-3.21	-0.63	9.620	.975	-.94	2.75
Emotional Control	48.87	9.357	-1.814	-2.37	.10	9.903	.146	-1.76	2.04
Behavioural Regulation	47.93	8.612	-3.592**	-3.20	-0.93	8.289	-.700	-2.15	1.03
Initiate	48.64	9.010	-2.254*	-2.54	-.17	8.678	-.735	-2.28	1.05
Working Memory	48.85	8.893	-1.931	-2.32	.02	8.617	-2.266*	-3.54	-.24
Plan/Organise	49.20	9.083	-1.308	-2.00	.40	9.242	-.910	-2.58	.96
Organisation of Materials	50.76	9.669	1.175	-.51	2.03	7.222	-.817	-1.95	.81
Monitor	46.38	9.686	-5.594**	-4.90	-2.35	8.576	-1.155	-2.61	.69
Metacognition	48.45	8.744	-2.633**	-2.70	-.39	8.089	-1.489	-2.73	.39
Global Executive Composite	48.42	8.684	-2.705**	-2.73	-.43	8.096	-1.464	-2.71	.41

Note: df = 223 for PF, df = 105 for TF;

* Significant at $p < .05$ level, **Significant at $p < .01$ level

Evaluation of the NEPSY-II subtests. Performance on the NEPSY-II Social Perception tasks (ToM and AR) did not differ significantly between the children in our sample and those in the American sample ($d = .01 - .06$). On other measures of EF and attention however, performance was varied. Children in our sample performed significantly lower on the Animal Sorting task ($p < .001$, $d = .58$) and significantly higher on Inhibition - Switching ($p < .001$; $d = .45$). On the slightly easier Inhibition task however, the difference between our sample and the standardisation sample was not significant. Our sample mean was also significantly higher than that of the standardisation sample on the Inhibition – Naming task ($p = .001$; $d = .34$). Mean performance on NEPSY-II subtests compared to the U.S mean of 10 is shown in Table 8 and depicted in Figure 5.

Table 8

One-sample t-tests comparing NEPSY-II scores from NZ sample with US test mean of 10 (SD = 3).

	NEPSY-II				
	Mean	SD	<i>t</i>	95% CI	
AR	9.82	2.675	-.723	-.67	.31
ToM	9.96	2.553	-.180	-.51	.42
Animal Sorting	8.21	3.182	-5.297**	-2.46	-1.12
Naming	11.04	3.180	3.561**	.46	1.62
Inhibition	10.54	3.159	1.865	-.03	1.12
Switching	11.26	2.615	4.525**	.71	1.82

Note: $df = 117$ except for Animal Sorting ($df = 88$) and Switching ($df = 87$);

**Significant at $p < .01$ level

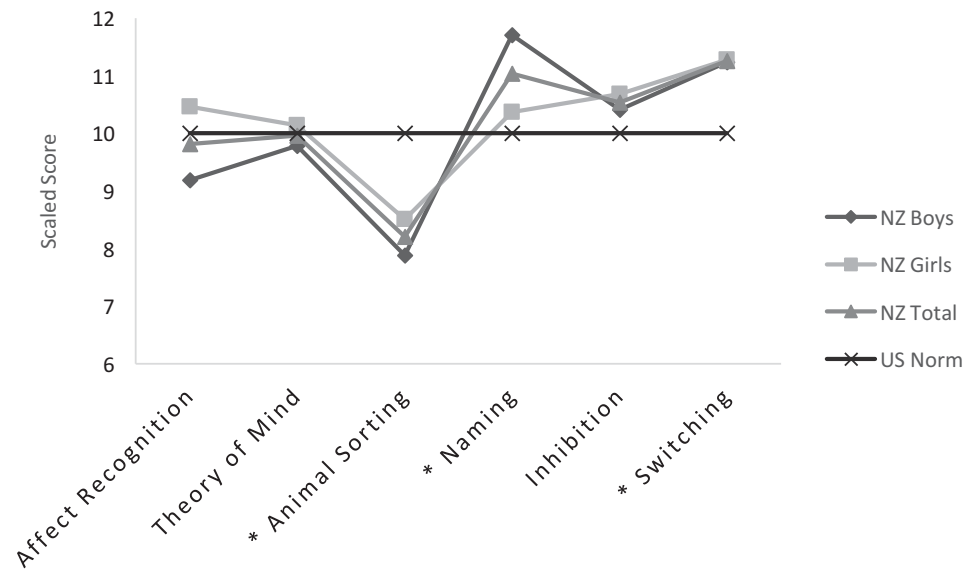


Figure 5. NEPSY-II means by gender and total sample, compared to the U.S. norm.

*Total significant at $p < .01$ level

Discussion. The results of this study indicate that there are significant differences between the scores of our NZ sample and the American standardisation sample particularly on the BASC-2, with a tendency for NZ parents and teachers to report fewer problem behaviours and more adaptive behaviours than the standardised means presented in the BASC-2 manual. In particular, teachers in our sample appear to rate children significantly higher on adaptive behaviours and significantly lower on problem behaviours than their American counterparts. The relationship between parent and teacher ratings though will be further discussed in Study Two. Differences were also noted on some domains of the BRIEF, however the effect sizes were small, indicating that this finding may not be replicable in future research.

The significant differences between the scores of our sample and the American standardisation sample suggest that the norms for the BASC-2, and to some extent the BRIEF, may not be appropriate for a NZ sample. There are a number of reasons why this might be the case. It may simply be that NZ children exhibit a lower prevalence of problem behaviours than the U.S, but other possibilities are that the prevalence of problem behaviours is similar, but these behaviours are masked by other factors. For example, Reynolds and Kaumphaus (2004) note that the standardisation sample used to develop the BASC-2 norms reported a higher proportion of behaviour problems than average. Thus, what appears to be differences across samples may be the result of skewed norms⁶.

Other possibilities include that discrepancies may be due to cultural differences in reporting (e.g., Ministry of Health (MOH), 2012), differences in the way the individual items are understood cross-culturally, cultural differences in behaviour itself, or a different cultural conceptualisation of what constitutes an emotional and behavioural problem. Therefore, items would be less relevant for NZ children. It may also be the case that NZ teachers have a more relaxed teaching style than U.S teachers, or any other number of differences in the teaching climate across cultures. Study-related factors such as the sample size and socio-economic status of our sample may also affect results, however limitations and implications relating to all three studies will be discussed in detail in the conclusion section of this thesis. Further research is required to fully understand what may be underlying these results.

⁶ As mentioned in an earlier footnote, the BASC-3 has been recently released. The standardisation sample for the BASC-3 is reported to be more representative of the general population than its predecessor, and therefore future research using the BASC-3 may provide more insight as to whether there is a true difference in behaviour/reporting across cultures.

The mean for NZ children on the BASC-2 is lower than that of the standardisation sample, which means that there is a larger gap between the “norm” and what is considered to be a clinically significant problem behaviour. This suggests that for a NZ child assessed using the BASC-2, their behaviour has to deviate from that of their peers even more so than in other countries in order for that behaviour to be considered clinically significant. This also suggests that the test is not as sensitive to clinically significant elevations in emotional and behavioural problems and that children who are experiencing significant problems in a domain may be under-diagnosed, which increases the risk of inaccurate treatment.

An interesting finding of our study was that 27 individuals (11.2%) scored in the clinically significant range in one or more domains of the BASC-2 PRS. This is higher than previous estimates of emotional and behavioural problems in NZ, with the New Zealand health survey annual update for 2012/2013 reporting that approximately 4% of NZ children had been identified as having emotional or behavioural problems (MOH, 2012). This report does state however that prevalence rates may be higher than recorded due to a tendency for people to under-report negative behaviours (MOH, 2012). The result in the current study suggests that in spite of reporting fewer problem behaviours than the U.S, there is still a reasonably high proportion of children presenting with clinically significant scores.

While there was no difference across cultures in performance on Social Perception tasks, results of this study indicate that EF performance on the NEPSY-II does differ between the NZ and American samples. These findings are consistent with researchers such as Sabbagh and colleagues (2006) who found cross-cultural differences in EF performance but not ToM performance in preschool children. Further research is required to investigate this topic more fully. The finding that children in our sample

scored lower on the animal sorting task may however be explained in some part by order effects. Carlson and Moses (2001) suggest that children should be administered tasks in the same order when conducting neuropsychological research, however as this was the last task that children were administered, it is possible that this result reflected participants reaching the end of their concentration after an hour of testing (as suggested by Gardiner, Hutchison, Müller, Kerns & Iarocci, 2017). Future research should consider manipulating the order in which children are administered tasks to determine if this is a true difference in performance or not.

Part Three

Overview. Part Three involved evaluation of scores as a function of gender and decile, as both may play a role in differences in behaviour across cultures. Independent-sample *t*-tests were used for these analyses and it was expected that there would be some differences in scores based on these variables.

Results

Gender comparisons. An Independent samples test indicated some significant differences by gender on the BASC-2 and performance measures. On the BASC-2 PRS, girls were on average rated higher on Functional Communication ($t = -2.621$; $p = .009$; $d = .36$) and the Adaptive Skills Composite ($t = -1.998$; $p = .047$; $d = .27$). On the TRS girls were rated higher on a scale of Leadership ($t = -2.263$; $p = .026$; $d = .44$) and boys tended to be rated higher on scales of Withdrawal ($t = 2.348$; $p = .021$; $d = .45$) and Learning Problems ($t = 2.300$; $p = .024$; $d = .44$).

On the performance measures, girls tended to perform better than boys on an AR task ($t = -2.646$; $p = .009$; $d = .49$); and boys better than girls on a timed naming task (t

= 2.330; $p = .022$; $d = .43$). Descriptive statistics for boys, girls and the total sample are depicted in Table 9.

Table 9
Descriptive statistics by gender

		Composite scores																	
		AR			ToM			Animal Sorting			Naming			Inhibition			Switching		
		Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
NEPSY-II																			
Performance (n = 118)	Mean (s.d.) Range	9.19 (2.96) 2-14	10.46 (2.21) 4-15	9.82 (2.68) 2-15	9.78 (2.72) 3-15	10.14 (2.39) 3-14	9.96 (2.55) 3-15	7.88 (3.44) 2-17	8.51 (2.94) 2-18	8.21 (3.18) 2-18	11.71 (3.04) 2-16	10.37 (3.20) 3-15	11.04 (3.18) 2-16	10.41 (3.20) 3-17	10.68 (3.14) 2-16	10.54 (3.16) 2-17	11.24 (2.66) 5-16	11.28 (2.60) 7-17	11.28 (2.60) 7-17
BASC-2																			
Parent (n = 216)	Mean (s.d.) Range	46.80 (7.88) 34-68	46.56 (7.23) 34-70	46.68 (7.54) 34-70	46.18 (9.15) 31-77	45.90 (8.72) 32-76	46.04 (8.91) 31-77	47.01 (7.41) 33-66	46.60 (7.21) 34-66	46.80 (7.29) 33-66	-	-	-	51.46 (7.82) 36-71	53.53 (7.43) 37-68	52.50 (7.68) 36-71	55.07 (7.60) 37-68	57.18 (7.09) 43-71	56.08 (7.40) 37-71
Teacher (n = 106)	Mean (s.d.) Range	45.05 (6.34) 40-64	45.04 (4.54) 41-60	45.05 (5.53) 40-64	49.16 (8.02) 37-72	48.32 (11.57) 38-94	49.07 (10.56) 38-94	46.16 (7.95) 37-72	44.76 (6.16) 38-75	45.50 (7.20) 37-75	44.95 (8.51) 35-71	42.60 (4.73) 38-54	43.83 (7.04) 35-71	55.07 (7.60) 37-68	57.18 (7.09) 43-71	56.08 (7.40) 37-71	55.07 (7.60) 37-68	57.18 (7.09) 43-71	56.08 (7.40) 37-71
BRIEF																			
Parent (n = 222)	Mean (s.d.) Range	48.07 (8.62) 34-71	47.79 (8.65) 35-79	47.93 (8.61) 34-79	48.31 (8.51) 33-67	48.60 (9.02) 32-77	48.45 (8.74) 32-77	48.52 (8.63) 34-78	48.33 (8.78) 34-78	48.42 (8.68) 34-78	48.50 (7.20) 37-75	42.60 (4.73) 38-54	43.83 (7.04) 35-71	49.91 (9.17) 37-82	47.62 (6.55) 40-66	48.83 (8.09) 37-82	49.64 (8.93) 38-80	47.96 (7.03) 40-74	48.85 (8.10) 38-80
Teacher (n = 106)	Mean (s.d.) Range	49.56 (8.98) 41-83	49.30 (7.51) 42-86	49.44 (8.29) 41-86	49.91 (9.17) 37-82	47.62 (6.55) 40-66	48.83 (8.09) 37-82	49.64 (8.93) 38-80	47.96 (7.03) 40-74	48.85 (8.10) 38-80	48.50 (7.20) 37-75	42.60 (4.73) 38-54	43.83 (7.04) 35-71	49.91 (9.17) 37-82	47.62 (6.55) 40-66	48.83 (8.09) 37-82	49.64 (8.93) 38-80	47.96 (7.03) 40-74	48.85 (8.10) 38-80

Comparisons across decile. Correlations are depicted in Table 10. The correlation between AR and decile was positive but not significant ($r = .14; p = \text{n.s.}$)⁷. Significant correlations were found between decile and parent-rated Metacognition ($r = -.17; p = .010$ ⁸) and with the Global Executive Composite ($r = -.17; p = .012$), suggesting that children from higher decile schools may have more developed executive skills. Further investigation showed that on the BRIEF PF domains, decile was significantly correlated with Initiation ($r = -.16; p = .014$), Working Memory ($r = -.16; p = .020$) and Planning/Organisation ($r = -.22; p = .001$) domains. On the performance measures of EF (NEPSY-II subtests) there were no significant correlations between decile and performance.

⁷ Higher decile associated with higher SES

⁸ Negative correlations in these comparisons should be interpreted as an increase in positive behaviours as a high score on the BRIEF, and on the BASC-2 problem domains, indicates difficulties in that domain.

Table 10.

Correlations between decile and test scores

NEPSY-II		Decile
Performance (n = 118)	AR	.14
	ToM	.16
	AS	.07
	Naming	-.03
	Inhibition	-.05
	Switching	.02
BASC-2		
Parent (n = 216)	Ext. Probs	-.03
	Int. Probs	-.08
	BSI	-.09
	Adaptive skills	.14*
Teacher (n = 106)	Ext. Probs	.01
	Int. Probs	.00
	BSI	.04
	Adaptive Skills	-.05
BRIEF		
Parent (n = 222)	BRI	-.08
	MI	-.17**
	GEC	-.17*
Teacher (n = 106)	BRI	-.10
	MI	-.07
	GEC	-.09

Note. *Significant at $p < .05$ level; **Significant at $p < .01$ level

Discussion

Gender comparisons. Comparisons across gender revealed that there were some differences both in performance and in the way they were rated by parents and teachers, with scores reflecting stereotypical patterns. Girls were rated as higher on positive attributes such as Functional Communication, Adaptive Skills, and Leadership, whereas boys seemed to be rated higher on problem behaviours (Withdrawal and Learning Problems). On all other BASC-2 and BRIEF domains there were no significant differences across gender.

On performance tasks, scores also differed as a function of gender, with girls performing better on the AR subtest and boys performing better on the timed Naming task. Our sample was not large or diverse enough to make comparisons based on ethnicity.

Concerns have been raised in previous research (e.g., Clay et al., 2008) that responses on rating scales of emotion and behaviour may differ as a function of the gender of the rater and/or the child, however analyses comparing this could not be conducted due to a very low number of male responders ($n = 11$). It may be that fathers were less likely to participate in the study due to the lengthy pesticide questionnaire which would have involved a large time commitment for the larger study. Within participating schools, it may also be that mothers were more involved with their child's schooling or health-related activities. It may be that future research conducted in NZ with a larger sample could investigate differences in responding across gender.

Comparisons across decile. Decile did appear to affect scores, with children from higher socio-economic areas being rated more positively on domains of

metacognition and adaptive skills by parents. Teacher ratings did not appear to change as a function of decile.

There may be a number of reasons why parents ratings appear to change as a function of decile. First, emotional and behavioural disorders are more prevalent in lower income areas (MOH, 2012). Second, expectations from parents may differ according to SES and parental education, leading to differences in the way parents interpret 'problem' behaviours. And third, children in higher income areas have likely had greater access to resources and positive relational influences which may lead to relative differences in behaviour compared to children from lower income areas.

CHAPTER EIGHT

Study Two

Overview

Previous research has indicated that parents and teachers typically have low to moderate concordance on behaviour rating scales, potentially due to differences in behaviour across contexts (Clay et al., 2008; Achenbach et al., 1987). This study seeks to extend previous research by evaluating the relationship between parent and teacher ratings of functioning in a NZ sample of typically developing children.

Correlations and paired samples *t*-tests between parent and teacher informant reports will be examined, both on the BASC-2 and the BRIEF. Based on the literature it is expected that the correlation between parent and teacher reports would be low to moderate.

Results

Relationship between parent and teacher responses on the BASC-2. To determine the relationship between parent and teacher responses on the BASC-2, bivariate correlations were conducted. As can be seen in Table 11, all correlations on matched domains were significantly correlated at either the $p < .01$ or $p < .05$ level except for the domains of Atypicality, Adaptability, and Functional Communication. Significant Pearson's *r* values ranged from $r = .22$ ($p = .026$) on the Social Skills domain to $r = .49$ ($p < .001$) on the domain of Withdrawal. This provides evidence for a low-moderate relationship between parent and teacher responses on the BASC-2.

Interestingly many of the domains of the PRS were more strongly correlated with a *different* domain on the TRS than the domain of the same name. Similarly, many of the domains on the TRS were more strongly correlated with a different domain on the

PRS than the matched domain. For example, parent-rated aggression was more strongly correlated with teacher-rated hyperactivity ($r = .33$; $p = .001$) than teacher-rated aggression ($r = .23$; $p = .025$), though this difference was not significant.

Paired samples *t*-tests for parent and teacher ratings on the BASC-2 domains are shown in Table 12. Teachers ratings on the adaptive skills domain were significantly higher than those of parents, with a moderate effect size of $d = .52$ ($p = .016$). Parents and teachers also differed significantly on domains of externalising problems ($p = .001$; $d = .34$) and behavioural symptoms ($p = .009$; $d = .32$), with parents tending to report more problem behaviours than teachers. Internalising problem scores did not differ significantly as a function of rater.

Table 11
Correlation between parent and teacher responses on the BASC-2

BASC-2 PRS																	
	Hyp.	Agg.	Cond. Probs.	Ext. Probs	Anx.	Dep.	Som.	Int. Probs	Atyp.	Withdrawal	Attn. probs.	BSI	Adapt.	Social Skills	Leadership	Func. Comm.	Adapt. skills
BASC-2 TRS																	
Hyperactivity	.32**	.33**	.30**	.37**	-.11	-.17	-.08	-.15	.05	-.14	.31**	.15	-.12	-.07	.05	-.20*	-.14
Aggression	.18	.23*	.27**	.26**	-.01	-.05	-.02	-.04	.10	.09	.29**	.20*	-.19	-.22*	-.10	-.28**	-.26*
Conduct problems	.22*	.24*	.30**	.29**	-.12	-.14	-.04	-.13	.03	.01	.27**	.15	-.11	-.15	-.02	-.25*	-.18
<i>Externalising problems</i>	.27**	.29**	.32**	.34**	-.10	-.14	-.07	-.13	.06	-.03	.31**	.17	-.15	-.16	-.02	-.25*	-.20*
Anxiety	-.13	-.06	-.09	-.11	.35**	.25*	.19	.33**	.03	.34**	.08	.17	-.26*	-.09	-.19	-.08	-.14
Depression	-.14	.09	.02	-.01	.27**	.33**	.21*	.33**	.01	.37**	.07	.22*	-.22*	-.18	-.25*	-.16	-.22*
Somatisation	-.05	.00	-.06	-.04	.17	.09	.27**	.23*	-.10	.09	-.03	.04	-.11	-.00	.03	.05	-.01
<i>Internalising problems</i>	-.12	.03	-.03	-.05	.33**	.30**	.28**	.38**	.00	.34**	.08	.21*	-.25*	-.12	-.19	-.10	-.17
Atypicality	.06	.13	.25*	.17	.10	.12	.14	.14	.07	.19	.22*	.21*	-.12	-.12	-.18	-.18	-.17
Withdrawal	-.06	.04	.01	.00	.27**	.26**	.16	.29**	.17	.49**	.07	.29**	-.16	-.24*	-.32**	-.20*	-.24*
Attention problems	.22*	.21*	.24*	.25*	-.01	-.03	.08	.02	.00	-.05	.31**	.15	-.12	-.16	-.04	-.23*	-.21*
<i>Behavioural Symptoms Index</i>	.10	.20*	.21*	.20*	.13	.13	.12	.16	.08	.23*	.25*	.26**	-.20*	-.20*	-.19	-.26**	-.26**
Adaptability	-.00	-.08	-.02	-.04	-.12	-.15	-.08	-.14	-.03	-.22*	-.17	-.17	.15	.15	.08	.15	.17
Social Skills	-.02	-.09	-.13	-.09	.06	-.01	-.01	.02	.06	-.15	-.20	-.10	.10	.22*	.07	.17	.18
Leadership	.06	-.06	-.06	-.02	-.20*	-.27**	-.18	-.27**	-.06	-.40**	-.15	-.24*	.18	.29**	.31**	.23*	.29**
Functional communication	-.00	-.05	-.06	-.04	-.14	-.26**	-.10	-.20*	.02	-.27**	-.12	-.17	.16	.17	.25*	.15	.22*
<i>Adaptive skills composite</i>	-.01	-.10	-.10	-.08	-.10	-.17	-.12	-.16	-.01	-.27**	-.20*	-.19	.16	.22*	.19	.22*	.24*

Note. Matched correlations are in bold. High scores on the Externalising Problems Composite, Internalising Problems Composite, and Behavioural Symptoms Index indicate a high number of problem behaviours, whereas high scores on the Adaptive Skills Composite indicate a high number of positive behaviours. Thus, negative correlations between them are to be expected. *Significant at the $p < .05$ level **Significant at the $p < .01$ level. $N = 99$

Table 12

Paired samples t-tests for BASC-2 composites across parent and teacher raters.

	Parent Mean (s.d.)	Teacher Mean (s.d.)	<i>t</i>	Cohen's <i>d</i>	95% CI	
Externalising problems	47.48 (7.87)	45.17 (5.62)	2.89	.34**	.724	3.902
Internalising problems	47.77 (9.57)	48.72 (10.48)	-.85	.09	-3.205	1.286
Behavioural Symptoms Index	48.02 (7.39)	45.63 (7.35)	2.66	.32**	.607	4.181
Adaptive skills	52.13 (7.69)	56.08 (7.60)	-4.17	.52**	-5.827	-2.072

Note. $N = 99$ for all domains except internalising which had $n = 98$. ** Significant at the $p < .01$ level.

Relationship between parent and teacher responses on the BRIEF. Bivariate correlations were conducted to determine the relationship between parent and teacher responses on the BRIEF. As indicated in Table 13, there was a significant correlation between parent and teacher responses on most matched domains with the exception of Shift and Emotional Control which were not significantly correlated. Interestingly, all but two (Monitor and Inhibit) teacher-rated domains correlated more strongly with the plan/organise domain on the BRIEF PF than with the matched parent-rated domain. A paired samples *t*-test, depicted in Table 14, indicated that there were no significant differences in composite scores between parent and teacher ratings on the BRIEF.

Table 13

Correlation between parent and teacher responses on the BRIEF

BRIEF PF											
	Inhibit	Shift	Emotional Control	Behavioural Regulation Index	Initiate	Working memory	Plan/Organise	Organisation of Materials	Monitor	Metacognition Index	Global executive composite
Inhibit	.32**	-.05	.02	.13	-.01	.17	.30**	.10	.33**	.27**	.21*
Shift	.12	.09	.16	.15	.16	.16	.34**	.21*	.18	.27**	.21*
Emotional Control	.06	-.01	.10	.07	.08	.05	.24*	.20*	.13	.19	.13
Behavioural Regulation Index	.19*	.01	.12	.14	.10	.15	.34**	.20*	.26**	.28**	.22*
Initiate	.10	.02	.13	.10	.21*	.20*	.36**	.12	.18	.29**	.20*
Working memory	.07	-.08	.05	.02	.19*	.27**	.40**	.18	.28**	.35**	.22*
Plan/Organise	.14	-.09	.06	.05	.18	.22*	.40**	.25*	.25**	.35**	.22*
Organisation of Materials	.11	-.01	.11	.09	.19	.33**	.33**	.21*	.34**	.37**	.28**
Monitor	.26**	-.11	.04	.09	.10	.18	.36**	.25*	.39**	.35**	.25*
Metacognition Index	.14	-.09	.07	.06	.18	.26**	.41**	.22*	.32**	.37**	.25*
Global executive composite	.17	-.05	.10	.10	.15	.22*	.40**	.21*	.33**	.36**	.25*

*Significant at the $p < .05$ level **Significant at the $p < .01$ level. $N = 104$

Table 14

Paired samples t-tests for BRIEF composites across parent and teacher raters.

	Parent Mean (s.d.)	Teacher Mean (s.d.)	<i>t</i>	Cohen's <i>d</i>	95% CI	
Behavioural Regulation Index	48.70 (9.22)	49.49 (8.31)	-.71	.09	-3.010	1.425
Metacognition Index	49.09 (8.78)	48.88 (8.14)	.23	.02	-1.635	2.058
Global Executive Composite	49.38 (9.14)	48.89 (8.14)	.46	.06	-1.584	2.545

*Significant at the $p < .05$ level ** Significant at the $p < .01$ level.

Note. $N = 106$

Discussion

Relationship between parent and teacher responses on the BASC-2. The results of this study suggest that, as hypothesised, there was a low-moderate concordance between parent and teacher ratings on the BASC-2. Interestingly, while there was a significant correlation across all matched domains (e.g., Aggression-Aggression), in a number of instances there was a stronger correlation with a non-matched domain (e.g., Hyperactivity-Aggression). The finding that many of the domains on the PRS were more strongly correlated with a domain *other than* the matched domain on the TRS (and vice versa) suggests that perhaps parents and teachers are rating the same behaviours in different ways as a result of the context in which they are observing the child. For example, children may be more likely to be rated as hyperactive within a classroom setting, yet the same behaviours may be interpreted as aggression in the home.

An analysis of the composite scores revealed that on the Externalising Problems Composite and Behavioural Symptoms Index there were significant differences between

parent and teacher responses with parents reporting more problem behaviours than teachers. There were no significant differences on ratings of internalising problems. On the Adaptive Skills Composite however, teachers were significantly more likely to rate positive adaptive skills than parents. Interestingly though, both parents and teachers rated fewer problem behaviours and more adaptive behaviours than the American standardisation sample. These findings suggests that when comparing to the BASC-2 standardisation sample, parents and teachers tend to report more positive behaviours than their American counterparts, but this difference is larger when evaluating teacher responses.

Relationship between parent and teacher responses on the BRIEF. Similarly on the BRIEF, as hypothesised, there was a low-moderate correlation between parent and teacher responses on most matched domains. Unexpected though was that again, the majority of teacher-rated behaviours correlated with a different domain on the parent-form. In terms of composite scores, there were no significant differences between parent and teacher ratings. Taking specific domains and composite scores together, these findings indicate that ratings on the BRIEF do not differ significantly between parents and teachers.

CHAPTER NINE

Study Three

Overview

This study will report how informant ratings relate to the child's performance on objective measures of functioning. Correlations between parent and teacher reports with the child's performance on Social Perception and EF tasks will be analysed. Research evaluating the relationship between subjective and objective measures has yielded mixed results thus the current study is an exploratory investigation of the relationship.

Results

Evaluation of BASC-2 ratings with NEPSY-II performance. As can be seen in Table 15, there was minimal relation between the BASC-2 domains (parent and teacher subjective ratings) and the child's performance on the NEPSY-II subtests. However, Adaptive Skills as rated by parents were positively correlated with performance on the ToM task ($r = .25$; $p = .013$), and Adaptive Skills as rated by teachers were positively correlated with performance on an AR task ($r = .27$; $p = .009$).

Analyses of the BASC-2 Content Scales revealed Teacher-rated Emotional Self-Control to be positively correlated with performance on a cognitive switching task ($r = .36$; $p = .028$). Otherwise, neither the EF Content Scale nor the Emotional Self-Control Content Scale was related to performance on Social Perception tasks. In particular, the EF Content Scale was not significantly related to EF performance.

Table 15

Correlations of BASC-2 composite and content scales with NEPSY-II scores

	AR (n)	ToM (n)	Animal Sorting (n)	Naming (n)	Inhibition (n)	Switching (n)
BASC-2						
Composites						
Externalising Problems						
Parent	-.05 (97)	-.12 (97)	.03 (76)	-.19 (91)	-.08 (97)	-.07 (75)
Teacher	.00 (94)	-.02 (94)	.03 (74)	-.16 (94)	.04 (94)	.19 (73)
Internalising Problems						
Parent	-.10 (97)	-.08 (97)	-.06 (76)	-.16 (97)	-.15 (97)	-.26* (75)
Teacher	-.11 (93)	.03 (93)	-.14 (73)	-.01 (93)	.08 (93)	.07 (72)
BSI ^a						
Parent	-.14 (97)	-.19 (97)	.02 (76)	-.20 (97)	-.19 (97)	-.21 (75)
Teacher	-.13 (94)	-.03 (94)	-.03 (74)	-.16 (94)	.02 (94)	.17 (73)
Adaptive Skills						
Parent	.15 (97)	.25* (97)	.02 (76)	.14 (97)	.12 (97)	.06 (75)
Teacher	.27** (94)	.20 (94)	-.14 (74)	.19 (94)	.08 (94)	-.10 (73)
BASC-2						
Content scales						
Emotional Self-control						
Parent	-.08 (40)	-.21 (40)	.25 (32)	-.22 (40)	-.10 (40)	-.02 (32)
Teacher	-.18 (55)	-.22 (55)	.14 (39)	-.21 (55)	-.08 (55)	.36* (38)
Executive Functioning						
Parent	-.14 (40)	-.22 (40)	.12 (32)	-.21 (40)	-.05 (40)	.03 (32)
Teacher	.10 (55)	-.15 (55)	.06 (39)	-.13 (55)	-.04 (55)	.26 (38)

*Significant at the $p < .05$ level ** Significant at the $p < .01$ level. ^aBSI = Behavioural Symptoms Index.

Evaluation of BRIEF ratings with NEPSY-II performance. Correlations between parent and teacher ratings on the BRIEF and the child's performance on the NEPSY-II subtests are depicted in Table 16. The few significant correlations of parent and teacher ratings on the BRIEF with the child's performance on the NEPSY-II subtests were all negative, i.e., more reported difficulties on the BRIEF was associated with lower performance on selected NEPSY-II subtests. Parent-rated scores on the Metacognition and Global Executive Composites were significantly correlated with both Social Perception tasks while teacher-rated scores were significantly correlated with performance on the Naming task. The Behaviour Regulation Index was not correlated with any of the NEPSY-II subtests on either the parent or teacher forms.

Table 16

Correlations of BRIEF composites with NEPSY-II scores

	NEPSY-II					
	AR (n)	ToM (n)	Animal Sorting (n)	Naming (n)	Inhibition (n)	Switching (n)
BRIEF Composites						
BRI						
Parent	-.19 (102)	-.15 (102)	.01 (78)	-.17 (102)	-.18 (102)	-.20 (77)
Teacher	-.15 (94)	.01 (94)	-.01 (74)	-.15 (94)	-.06 (94)	.14 (73)
MI						
Parent	-.21* (101)	-.22* (101)	.10 (77)	-.16 (101)	-.21* (101)	-.12 (76)
Teacher	-.19 (93)	-.12 (93)	.06 (73)	-.26* (93)	-.13 (93)	.12 (73)
GEC						
Parent	-.26** (101)	-.24** (101)	.06 (77)	-.15 (101)	-.19 (101)	-.10 (76)
Teacher	-.19 (93)	-.09 (93)	-.01 (73)	-.25* (93)	-.13 (93)	.11 (73)

Note. BRI = Behaviour Regulation Index; MI = Metacognition Index; GEC = Global Executive Composite

Discussion

Evaluation of BASC-2 ratings with NEPSY-II performance. This study found minimal relationship between behavioural ratings on the BASC-2 and NEPSY-II EF and ToM performance. More broadly, behaviours related to EF on this broadband *subjective* assessment tool were not strongly related to performance on an *objective* measure of EF ability.

Consistent with research that suggests that Social Perception is associated with prosocial behaviours (Slaughter et al., 2002) the current study also found that adaptive

skills as rated by parents and teachers on the BASC-2 were related to the child's performance on Social Perception tasks.

The finding that parent-rated internalising problems and performance on a switching task were negatively correlated is consistent with the large body of research that suggests cognitive biases underlie internalising problems such as anxiety and depression (Gotlib & Joormann, 2010). In fact, a meta-analysis by Wagner and colleagues (2015) found a standard mean difference of .444 ($p < .01$) when comparing children diagnosed with Major Depressive Disorder and controls on their switching ability, with other cognitive and EFs also showing significant differences across groups. Children with internalising problems may have difficulty with this task due to an impaired ability to see things a different way. In other words, they may have a limited capacity to hold two differing mental concepts in their head simultaneously and switch between them as required. In this task, this is expressed as a limited ability to remember and switch between two differing rules.

Finally, teacher-rated emotional self-control was related to performance on the Switching task. This finding makes sense, as the Switching task requires self-control to inhibit the automatic response. However, given that *emotional* self-control is typically considered a 'hot' EF process; it is interesting that it was positively correlated with Switching, a 'cooler', more cognitive-based task. It was also interesting that the EF as measured by the BASC-2 EF content scale was not related to any of the EF performance tasks suggesting that the BASC-2 may not be a good indicator of EF performance on the NEPSY-II.

Evaluation of BRIEF ratings with NEPSY-II performance. There were few significant relationships between subjective (narrowband) and objective measures of EF using the BRIEF and the NEPSY-II. This was consistent with Anderson (2002) who reported low-moderate correlations between the two. Such findings seem contradictory considering both are designed to measure EF but may be a result of subjective and objective measures of EF assessing different things (as suggested by Rubin & Pepler, in 2013).

Interestingly, the Behavioural Regulation index which was thought to be more closely aligned with the NEPSY-II domains (due particularly to previous research suggesting close links between Inhibition and Emotional Control with both Social Perception and EF), was not correlated with any of the NEPSY-II subtests on either the parent or teacher forms. This finding is consistent with earlier results that indicated no significant relationship between EF skills as determined by the BASC-2 rating scale and EF performance.

While parent ratings on the BRIEF appear to be related to performance on the Social Perception measures from the NEPSY-II (AR and ToM), teacher ratings appear to be related to a child's performance on a timed naming task. This would seem to suggest that parent and teacher ratings on the BRIEF reflect different aspects of EF. Again considering the role of 'hot' and 'cool' EFs these findings would seem to suggest that parents and teachers observe behaviours in different motivational contexts. In other words, parent ratings reflect the more 'hot' emotionally-driven EF behaviours associated with ToM and Affect Recognition, whereas teacher ratings are more reflective of cognitive 'cool' EFs that may be more appropriate for a classroom setting. For example, the Naming task requires a child to perform with speed and accuracy

which is much more likely to be important in a classroom setting than within the home environment. Further research is required to tease apart this relationship.

CHAPTER TEN

Conclusion

Summary of Findings

The results of the current thesis indicate that there is little relationship between subjective informant reports (BASC-2 and BRIEF) and the child's performance on the NEPSY-II EF and Social Perception subtests. This is consistent with Rubin and Pepler (2013) who suggested that subjective and objective measures of EF may actually assess different constructs. This study also extends previous research by providing preliminary evidence that the relationship between subjective and objective measures may differ by informant, with parents responses more highly correlated with affective Social Perception tasks and teacher responses more strongly correlated with cognitive EF tasks. Consistent with existing research, parent and teacher ratings of functioning were low-moderately correlated.

A unique feature of the research covered in this thesis was evaluation of the prevalence of BASC-2 Critical Items. These were more prevalent than expected given that this was a non-clinical sample and conveyed some interesting information about our population sample. For example, in our study 25 parents indicated that their child sometimes or often reported "I want to die", "I wish I were dead", or "I want to kill myself". Over a quarter of these parents identified that they were concerned about the behaviour, though most of these had already sought mental health involvement. Our study also found that there were higher levels of clinically significant externalising or internalising problems than those reported by the Ministry of Health (MOH, 2012). Taken together, the findings of the current study suggests that in spite of reporting fewer problem behaviours than American parents and teachers, there is still a reasonably high proportion of children presenting with clinically significant scores. This finding is

consistent with MOH (2012) reports that the estimation of emotional and behavioural disorders in NZ is likely higher than recorded due to a tendency for people to under-report negative behaviours.

Implications for Clinical Practice

These results, combined with those of Rubin and Pepler (2013) and other emerging studies in this area, have several implications for practitioners. First, when considering EF and ToM assessment clinicians should be wary of inferring behaviour from cognition and vice versa. Behavioural measures completed by parents and teachers provide differing information from each other and from child performance on cognitive tasks. Use of only one of these measures is likely to provide only part of the picture. For example, subjective ratings of EF behaviours on the BRIEF may suggest impaired EF which may not necessarily be reflected in assessment of EF performance. The discrepancy in responding between parents and teachers likely reflects the differences in behaviour across contexts and is informed by the rater's relationship and history with the child. In practice, utilising both parent and teacher reports allows for a more complete assessment of the child's functioning across contexts.

Second, when completing assessments with the BASC-2, Critical Items provide useful information, however consideration of age and social factors is also required. While death related items were more frequently reported in the general population than was expected, parents predominantly indicated that the behaviour was not of concern, suggesting that these items should not always be taken at face value, but rather as cues for further investigation. A small minority of parents expressed concern and yet had not sought out mental health services. Though the BASC-2 manual advises clinicians to enquire further about these items, no literature reviewed for this thesis indicated that

researchers had followed up these items when using the BASC-2. This study suggests that querying these items is important in research as well as in clinical practice.

Third, assessment of ToM is important to consider when assessing for executive deficits due to the high level of overlap between ToM and EF. The level of overlap is also likely to evolve as definitions and assessment tools for both constructs continue to develop. Finally, the results of the current thesis also indicate that caution should be used when applying norms for the BASC-2, BRIEF, and NEPSY-II to NZ school-aged children. The implications of these findings are that when comparing to the current standardisation sample, clinicians should be aware that NZ parents and especially teachers tend to rate children more positively on these measures, thus the measures may be less sensitive to impairment. This is especially true when using BASC-2 norms with NZ children, as the normative sample is reported to have a higher number of participants with behaviour problems than average. Therefore when compared to this group, NZ children may appear to be doing better than they actually are and prevalence of problem behaviours may be underestimated.

Limitations and Future Directions

Sample characteristics

Decile distribution of sample. One of the limitations of this research was that many of the schools that agreed to participate were higher decile schools, indicating that the income and SES of families attending the school is likely to be high. An implication of this may be that other factors (more qualified teachers, less financial stressors, higher parental education) may be impacting on how children behave and how their behaviour is interpreted by parents and teachers. Therefore findings may be more a reflection of the sample and less generalizable to lower decile schools. As data collection was time-

limited and our response rate for lower decile schools in the larger study was low, it was difficult to remedy this.

Recruitment efforts attempted to mitigate this problem early, over-sampling lower decile schools. However, the abundance of higher decile schools in our sample is due to a number of factors. First, schools were more likely to agree to participate if they were approached by a parent or teacher within the school. Contacts were utilised from people within CPHR who predominantly had connections with higher decile schools. Second, there were a number of other research studies being conducted at the same time, and many of the smaller, lower decile schools felt overwhelmed by the number of requests to participate and were already involved in other research. Third, some of the lower decile schools consisted of beginning or less experienced teachers and Principals were less inclined for them to be involved with additional research demands.

Previous research has suggested ToM to be associated with socioeconomic status (e.g., Cutting & Dunn, 1999; Liddle & Nettle, 2006). While we found no significant differences in ToM performance in the current study, it is possible that differences may have been observed with a more diverse sample.

Gender and ethnicity distribution of sample. In terms of gender, our sample was matched to the population. Unfortunately, the ethnic distribution of our sample was not and consisted mostly of NZ European students. Because the study was offered to such a large number of families, it seems that the response rate from other ethnicities was low. There are a number of reasons this might have been the case. First, the schools that chose to participate were not overly diverse. Second, parents may have self-selected out due to language barriers or the length of the questionnaire. For parents who had

English as a second language, the sheer volume of reading and writing involved in the larger study may have led them to self-select out of the study. When parent uptake of the study was discussed with school Principals, they informed study investigators that it was difficult even for the school to get parents to read and return school notices that were sent home. For the most part, Principals were unsurprised at the particular parents who opted out of the study.

In the initial development of the study, cultural advice was sought to ensure that the way in which the study was being conducted was culturally appropriate. However, there may have been a number of cultural reasons why people chose not to participate. For example, the third phase of the larger study involves the collection of urine and dust samples. It is possible that the collection and storage of urine samples may have concerned some cultural groups. As an example, NZ Maori view bodily fluids as items which need to be protected and handled properly, therefore consent to collect urine samples is not given lightly (Sporle & Koea, 2004). It may be that people opted out of participating as opposed to voicing any concerns directly.

The implication of having limited demographic variability in our sample is that we were unable to assess the extent to which ethnicity may contribute to performance on tasks. For instance, it has been previously suggested that a child's ability to perform appropriately on verbal ToM tasks may be culturally specific (Bosacki & Wilde Astington, 1999). It is possible then that with a more representative group there may be more variability in ToM performance. It is also unclear the extent to which ethnicity and cultural differences may be reflected in responses on the BASC and BRIEF questionnaires within our NZ sample. This may be an interesting avenue for future research to consider.

The measures

BASC and BRIEF questionnaires. At the time of writing, updates in both the BASC and BRIEF measures were released (BASC-3 and BRIEF-2). Had these updated measures been available earlier, they would have been used for the current research, however this was not an option when tests were selected and beginning to be administered. It is unclear to what extent the findings of the current research will be replicable with these updated measures.

NEPSY-II. The NEPSY-II was selected as it enables clinicians to assess a large range of cognitive functions with one tool based on a single normative sample as opposed to selecting a range of different instruments. A benefit of this is that variations in scores are less likely to be an artefact of comparing to different standardisation samples (Kemp, 2007). Unfortunately, as ToM development is relatively unstudied in middle childhood, it is unclear to what extent ToM is accurately captured by the NEPSY-II, and to what extent EF abilities impact on ToM performance. However, given the level of research with this age group at this point in time, the same could be said for any other measures of ToM performance.

Neuropsychological testing

Order effects. In the current study, test order was consistent across individuals, primarily to control for practice and interference effects within the assessment. One of the results in the current research was the finding that NZ children did not perform as well on the Animal Sorting task as their American counterparts. It is possible that the

reason for this may have been that the Animal Sorting task was always administered last. If this study were to be replicated in future, researchers should consider controlling for order effects by changing the order of administration for every second child tested. This alternate test order would also need to consider practice and interference effects.

Assessment period. Unfortunately a consequence of working with such a large study, and in collaborating with the various schools, parents and teachers involved in the study was that there was a substantial lag in the collection of data. In some cases there was a gap of approximately four to five months between when parent/teacher reports were received and the neuropsychological assessment with the child. This delay in testing, while unlikely to have drastically changed responses/performance, may still have impacted on the relationship between subjective and objective measures.

Testing environment. Typically a neuropsychological assessment would be completed in a standardised testing environment to eliminate extraneous factors that may impact on performance. Another unavoidable consequence of working within the school environment however, was that a standardised testing environment was rarely possible. Schools had limited space available and as a result assessments were frequently completed in storage rooms, staff rooms or resource rooms. Assessments were managed as best as possible given the conditions, though participants were still working within a stimulating, changeable environment with frequent disruptions. Participating children were used to working in this environment though and appeared to enjoy having time out of class and completing the assessments. For these reasons, issues of sustained attention were less of an issue than expected, and were noted by the

examiners when they occurred. Future research should consider completing the assessments outside of school hours in a standardised testing environment.

Conceptual Framework

The initial concept for this thesis had also involved developing a model of the relationship between EF and ToM. The more information reviewed and compiled for this thesis, the clearer it became that this was not possible with the data set for the current research. EF and ToM are both latent variables, theoretical constructs that are thought to be captured by assessment tools (i.e., subjective and objective measures). The key issue however, is that EF in particular is so loosely defined and encompasses numerous functions, that any assessment selected will capture a different function or series of functions.

This means then, that in order to develop a model that depicts the relationship between EF and ToM, one would need to be able to first be able to accurately differentiate between these highly linked constructs and second, be able to identify appropriate assessment tools to function as observed variables. What the current study, and researchers like Rubin and Pepler (2013) have identified however, is that subjective and objective measures appear to capture different aspects of these constructs; parent and teacher forms are not strongly correlated; and even in the NEPSY-II, domains of functioning such as Social Perception combine two latent variables that are theoretically linked (ToM and AR) but performance on these measures is not strongly correlated. Further, if Anderson (2002) is correct, there may be no measures of ToM available that do not also assess EF.

Results of the current study also indicate that norms for these measures are not necessarily appropriate for a NZ sample, thus using data from our sample might not be useful for developing a model of this relationship⁹. Further, statistical modelling requires data that is normally distributed. The distribution of scores in our sample was not. Although developing a model in the current study was not practical, future researchers may consider developing a model that evaluates the relationship between specific EFs and ToM. For instance Inhibition is thought to be one of the EFs most strongly related to preschool false belief understanding in ToM (Best et al., 2009). It would be interesting to know which EFs were most strongly related to ToM with this age group. The research would however have to consider how they might control for EF performance on ToM tasks.

Avenues for Future Research

Anderson (2002) commented that “longitudinal studies will be required to verify our understanding of EF development”. I concur with this statement, but would also stress the importance of including ToM development in these studies. An abundance of research is now being collated that shows these constructs to be closely linked. A longitudinal study would be able to consider social functioning as a whole and determine whether specific EFs or ToM are developing at particular ages/stages of development beyond preschool years and whether these developments are occurring simultaneously. It may be that this is the only way to determine whether EF and ToM are actually substantially different. Completing a longitudinal study of this nature would have implications for the assessment and treatment of conditions such as ADHD, ASD

⁹ It also appears that the BASC-2 norms were developed with a standardisation sample that included disproportionately high problem behaviours compared to the average, meaning it may not be a useful measure for evaluating this relationship. The BASC-3 may be more suitable for this analysis.

and Schizophrenia, especially as early interventions may be able to be developed that are targeted to address specific functions.

The findings of the current research indicate that the norms for the BASC-2, BRIEF, and NEPSY-II EF subtests are not necessarily suitable for assessment of NZ children. Future research is required to develop NZ norms for these measures or develop a measure within this cultural context to assess emotional, behavioural, and executive functioning of NZ children. In this way, the impact of SES, cultural background, and cultural differences in responding would be addressed and the measure would be more sensitive to clinical levels of impairment.

Overview of Author Reflections

While this section provides a brief outline, thorough reflections on the research process and the learning that occurred through conducting research are provided in Appendix D, including a discussion of how completion of this research impacted the author's professional development. In particular, being able to develop critical thinking skills, learning to draw together large volumes of information in a meaningful way, learning to monitor progress, reflect and respond to feedback are all skills that were vital for both research and clinical practice.

Through completing this research, the author was also able to develop increased awareness of and experience with assessment of cognition and behaviour, learn more about childhood cognitive development, and gain further insight into interpretation of neuropsychological test scores. Similarly learning how to respond flexibly and adapt to changing environments, learning to discuss risk effectively with parents and learning how to work with a team, with children and within the schooling environment were all

essential skills for clinical practice that were able to be developed throughout the research process. Completion of research and clinical placements also highlighted the importance of ongoing supervision and professional development to continue practicing and honing these skills.

Conclusions

Assessment of EF and ToM is complex and multifaceted. Broadband and narrowband measures are highly correlated but appear to provide substantially different information than that gained from objective assessment. Parent and teacher responses on subjective rating scales are not strongly related. Both parents and teachers appear to rate their children more positively than the parents and teachers in the American standardisation sample, yet there is still a high proportion of children with clinically significant scores. These findings have implications for clinicians and researchers working in this area, with further research indicated to address assessment of these functions for NZ children.

REFERENCES

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioural and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological bulletin*, *101*(2), 213.
- Achenbach, T. M., & Rescorla, L. (2001). ASEBA school-age forms & profiles.
- Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *8*(2), 71–82.
doi:10.1076/chin.8.2.71.8724
- Anderson, V., Northam, E., Hendy, J., & Wrennall, J. (2001). Pediatric neuropsychology: a clinical approach.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental psychology*, *45*(1), 190.
- Arnett, A. B., Pennington, B. F., Willcutt, E. G., DeFries, J. C., & Olson, R. K. (2015). Sex differences in ADHD symptom severity. *Journal of Child Psychology and Psychiatry*, *56*(6), 632-639.
- Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in psychology*, *5*.
- Banerjee, R., Watling, D., & Caputi, M. (2011). Peer relations and the understanding of faux pas: longitudinal evidence for bidirectional associations. *Child Development*, *82*(6), 1887–905. doi:10.1111/j.1467-8624.2011.01669.x

- Barker-Collo, S. L. (2003). Culture and validity of the Symptom Checklist-90-Revised and Profile of Mood States in a New Zealand student sample. *Cultural Diversity and Ethnic Minority Psychology, 9*(2), 185.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”?. *Cognition, 21*(1), 37-46.
- Baron-Cohen, S., & Swettenham, J. (1997). Theory of mind in autism : its relationship to executive function and central coherence. In *Handbook of Autism and Pervasive Developmental Disorders* (pp. 1–40).
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders, 34*(2), 163-175.
- Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: an integrative framework for the development of social skills. *Psychological Bulletin, 136*(1), 39–64.
doi:10.1037/a0017768
- Berk, L. E. (2006). *Child Development* (7th ed.). Boston: Pearson Education, Inc.
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental review, 29*(3), 180-200.
- Blakemore, S.-J., & Choudhury, S. (2006). Development of the adolescent brain: implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47*(3-4), 296–312.
doi:10.1111/j.1469-7610.2006.01611.x
- Bosacki, S., & Wilde Astington, J. (1999). Theory of Mind in preadolescence: Relations between social understanding and social competence. *Social Development, 8*(2), 237–255. doi:10.1111/1467-9507.00093

- Bosco, F. M., Gabbatore, I., & Tirassa, M. (2014). A broad assessment of theory of mind in adolescence: the complexity of mindreading. *Consciousness and Cognition, 24*, 84–97. doi:10.1016/j.concog.2014.01.003
- Brüne, M. (2005). Emotion recognition, “theory of mind,” and social behavior in schizophrenia. *Psychiatry Research, 133*(2–3), 135–147.
<https://doi.org/10.1016/j.psychres.2004.10.007>
- Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., ... Channon, S. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society: JINS, 12*(2), 194–209. doi:10.1017/S1355617706060310
- Caputi, M., Lecce, S., Pagnin, A., & Banerjee, R. (2012). Longitudinal effects of theory of mind on later peer relations: the role of prosocial behavior. *Developmental Psychology, 48*(1), 257–70.
<https://doi.org/10.1037/a0025402>
- Carlson, S. M., Koenig, M. A., & Harms, M. B. (2013). Theory of mind. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(4), 391-402.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child development, 72*(4), 1032-1053.
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*(2), 73–92.
doi:10.1002/icd.298
- Carter, A. S., Briggs-Gowan, M. J., & Davis, N. O. (2004). Assessment of young children’s social-emotional development and psychopathology: Recent

advances and recommendations for practice. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 45(1), 109–134.

- Clay, H., Surgenor, L. J., & Frampton, C. M. (2008). Assessing emotional and behavioural problems in children: Factors associated with multiple informant consistency in New Zealand. *New Zealand Journal of Psychology*, 37(1), 10–16.
- Collins, G. (2004). Small New Zealand primary schools: Current policy, its impact and some alternatives. *New Zealand Annual Review of Education*, 13, 63 - 78.
- Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: a systematic review from 1988 to 2010. *Clinical Psychology Review*, 31(5), 767–85.
doi:10.1016/j.cpr.2011.03.008
- David, D. P., Soeiro-De-Souza, M. G., Moreno, R. A., & Bio, D. S. (2014). Facial emotion recognition and its correlation with executive functions in bipolar patients and healthy controls. *Journal of Affective Disorders*, 152–154(1), 288–294. <https://doi.org/10.1016/j.jad.2013.09.027>
- Davis, J. L., & Matthews, R. N. (2010). NEPSY-II Review. *Journal of Psychoeducational Assessment*, 28(2), 175–182.
doi:10.1177/0734282909346716
- DeBellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, A., ... Boring, A. M. (2001). Sex Differences in Brain Maturation during Childhood and Adolescence. *Cerebral Cortex*, 11(6), 552–557.
- Demopoulos, C., Hopkins, J., & Davis, A. (2013). A comparison of social cognitive profiles in children with autism spectrum disorders and attention-deficit/hyperactivity disorder: A matter of quantitative but not qualitative

difference? *Journal of Autism and Developmental Disorders*, 43(5), 1157–1170. <https://doi.org/10.1007/s10803-012-1657y>

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development*, 84(3), 989-1003.

Dodell-Feder, D., DeLisi, L. E., & Hooker, C. I. (2014). The relationship between default mode network connectivity and social functioning in individuals at familial high-risk for schizophrenia. *Schizophrenia research*, 156(1), 87-95.

Education Counts. (2017). *School Rolls*. Wellington, NZ: Ministry of Education.

Retrieved from

<http://www.educationcounts.govt.nz/statistics/schooling/student-numbers/6028>

Education Review Office (2015). *Education Review Office*. Wellington, NZ: Education Review Office. Retrieved from <http://www.ero.govt.nz/review-reports/&Category=3>

Feigin, V. L., Theadom, A., Barker-Collo, S., Starkey, N. J., McPherson, K., Kahan, M., ... & Jones, K. (2013). Incidence of traumatic brain injury in New Zealand: a population-based study. *The Lancet Neurology*, 12(1), 53-64.

Fernando, K., Chard, L., Butcher, M., & McKay, C. (2003). Standardisation of the rey complex figure test in New Zealand children and adolescents. *New Zealand Journal of Psychology*, 32 (1), 33 - 38.

Frith, C. D., & Frith, U. (1999). Interacting minds - a biological basis. *Science*, 286(5445), 1692-1695.

Gardiner, E., Hutchison, S. M., Müller, U., Kerns, K. A., & Iarocci, G. (2017). Assessment of executive function in young children with and without ASD

using parent ratings and computerized tasks of executive function, *The Clinical Neuropsychologist*.

<http://dx.doi.org/10.1080/13854046.2017.1290139>

- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological bulletin*, *134*(1), 31.
- Gilmore, A., Croft, C., & Reid, N. (1981). Burt word reading test: New Zealand revision. *Wellington, New Zealand: New Zealand Council for Educational Research*.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000a). *Behavior rating inventory of executive function*. Odessa, FL: Psychological Assessment Resources.
- Gioia, G. a, Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000b). Behavior rating inventory of executive function. *Child Neuropsychology: A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *6*(3), 235–8. doi:10.1076/chin.6.3.235.3152
- Goldstein, S., Naglieri, J. A., Princiotta, D., & Otero, T. M. (2014). Introduction: A history of executive functioning as a theoretical and clinical construct. In S. Goldstein, & J. A. Naglieri. (Eds.), *Handbook of executive functioning* (pp. 3 – 12). New York: Springer.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance reality distinction. *Child development*, 26-37.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of child psychology and psychiatry*, *38*(5), 581-586.

- Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual review of clinical psychology*, 6, 285-312.
- Hale, J. B., & Fitzer, K. R. (2015). Evaluating orbital-ventral medial system regulation of personal attention: A critical need for neuropsychological assessment and intervention. *Applied Neuropsychology: Child*, 4(2), 106-115.
- Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, 20(3), 290–322.
<https://doi.org/10.1007/s11065-010-9138-6>
- Henry, J. D., Phillips, L. H., Crawford, J. R., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia*, 44(10), 1623–1628. <https://doi.org/10.1016/j.neuropsychologia.2006.03.020>
- Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., Zelazo, P. D., Wendy, S., Lee, C., & Lee, S. C. (2010). Developmental neuropsychology assessment of hot and cool executive function in young children : Age-related changes and individual differences. *Developmental Neuropsychology*, 28(October 2011), 37–41. doi:10.1207/s15326942dn2802
- Hughes, C. (2011). *Social understanding and social lives: From toddlerhood through to the transition to school*. Psychology Press.
- Hughes, C., & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions?. *Child and adolescent mental health*, 7(3), 131-142.

- Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social development, 13*(4), 590-619.
- Im-Bolter, N., Agostino, A., & Owens-Jaffray, K. (2016). Theory of mind in middle childhood and early adolescence: Different from before?. *Journal of experimental child psychology, 149*, 98-115.
- Isquith, P. K., Roth, R. M., & Gioia, G. (2013). Contribution of rating scales to the assessment of executive functions. *Applied Neuropsychology: Child, 2*(2), 125-132.
- Kemp, S. (2007). Overview. In *Essentials of NEPSY-II Assessment* (pp. 1–33).
- Kennerley, S. (2013). *Informant discrepancies in the assessment of Attention Deficit/Hyperactivity Disorder* (Doctoral dissertation, University of Otago).
- Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: A meta-analytic review. *Schizophrenia Bulletin, 36*(5), 1009–1019.
<https://doi.org/10.1093/schbul/sbn192>
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II*. San Antonio, TX: Pearson.
- Lai, M. C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(1), 11-24.
- Lenti, C., Giacobbe, A., & Pegna, C. (2000). Recognition of emotional facial expressions in depressed children and adolescents. *Perceptual and Motor Skills, 91*, 227–236.

- Liddle, B., & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*, 4(3), 231–244. doi:10.1556/JCEP.4.2006.3-4.3
- Mahy, C. E. V, Moses, L. J., & Pfeifer, J. H. (2014). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience*, 9, 68–81.
<https://doi.org/10.1016/j.dcn.2014.01.002>
- McCloskey, G., & Perkins, L. A. (2013). *Essentials of executive functions assessment*. Hoboken, NJ: Wiley.
- McClure, E. B., Pope, K., Hoberman, A. J., Pine, D. S., & Leibenluft, E. (2003). Facial expression recognition in adolescents with mood and anxiety disorders. *The American Journal of Psychiatry*, 160(6), 1172–1174.
<https://doi.org/10.1176/appi.ajp.160.6.1172>
- McNaughton, S., Phillips, G., & MacDonald, S. (2000). Curriculum channels and literacy development over the first year of instruction. *New Zealand Journal of Educational Studies*, 35.
- Ministry of Education. (2015). *School deciles*. NZ: New Zealand Ministry of Education. Retrieved from <http://www.education.govt.nz/school/running-a-school/resourcing/operational-funding/school-decile-ratings/>
- Ministry of Health. (2012). *New Zealand Health Survey Annual update of key findings 2012/13*. Wellington: Ministry of Health.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, 41(1), 49-100.

- New Zealand Psychologists Board. (2013). *Guidelines on the use of Psychometric Tests (February, 2013)*. NZ: New Zealand Psychologists Board.
- Ogden, J. A., & McFarlane-Nathan, G. (1997). Cultural bias in the neuropsychological assessment of young Maori men. *New Zealand Journal of Psychology, 26*, 2-12.
- Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*(9), 337–344. doi:10.1016/S1364-6613(99)01362-5
- Peterson, E., & Welsh, M. C. (2014). The development of hot and cool executive functions in childhood and adolescence: Are we getting warmer?. In *Handbook of executive functioning* (pp. 45-65). Springer New York.
- Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavior assessment system for children-Second edition*. Circle Pines, MN: American Guidance Services.
- Riccio, C. A., Sullivan, J. R., & Cohen, M. J. (2010). Introduction to Pediatric Neuropsychology. In *Neuropsychological assessment and intervention for childhood and adolescent disorders* (pp. 1–14). John Wiley & Sons.
- Robb, M. P., & Gillon, G. T. (2007). Speech rates of New Zealand English-and American Englishspeaking children. *Advances in Speech Language Pathology, 9*(2), 173-180.
- Roberts, J. R., & Carr, C. J. (2012). Policy statement: Pesticide exposure in children. *American Academy of Pediatrics 130* (6): e1757-1763.
- Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. *Applied neuropsychology, 12*(4), 190-201.

- Rosenqvist, J., Lahti-Nuutila, P., Urgesi, C., Holdnack, J., Kemp, S. L., & Laasonen, M. (2017). Neurocognitive functions in 3-to 15-year-old children: An international comparison. *Journal of the International Neuropsychological Society*, 23, 1-14. Doi: 10.1017/S1355617716001193
- Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). *Handbook of Executive Functioning*. (S. Goldstein & J. A. Naglieri, Eds.). New York, NY: Springer New York. doi:10.1007/978-1-4614-8106-5
- Rubin, K. H., & Pepler, D. J. (Eds.). (2013). *The development and treatment of childhood aggression*. Psychology Press.
- Sabbagh, M. a, Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind. A comparison of Chinese and U.S. preschoolers. *Psychological Science*, 17(1), 74–81. <https://doi.org/10.1111/j.1467-9280.2005.01667.x>
- Saracho, O. N. (2014). Theory of mind: children's understanding of mental states. *Early Child Development and Care*, 184(6), 949–961. doi:10.1080/03004430.2013.821985
- Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in cognitive sciences*, 19(2), 65-72.
- Sebastian, C., Fontaine, N., Bird, G., Blakemore, S., De Brito, S., McCrory, E., & Viding, E. (2012). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. doi:10.1093/scan/nsr023
- Semrud-Clikeman, M., & Ellsion, P. A. T. (2009). *Child Neuropsychology*. Springer US.
- Shahaeian, A., Nielsen, M., Peterson, C. C., & Slaughter, V. (2014). Cultural and family influences on children's theory of mind development: A comparison of

Australian and Iranian school-age children. *Journal of Cross-Cultural Psychology*, 45(4), 555-568.

- Shamay-Tsoory, S. G., Tibi-Elhanany, Y., & Aharon-Peretz, J. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Social Neuroscience*, 1(3-4), 149–166.
doi:10.1080/17470910600985589
- Silver, H., Goodman, C., Knoll, G., & Isakov, V. (2004). Brief emotion training improves recognition of facial emotions in chronic schizophrenia. A pilot study. *Psychiatry Research*, 128(2), 147–154.
<https://doi.org/10.1016/j.psychres.2004.06.002>
- Simonian, S. J., Beidel, D. C., Turner, S. M., Berkes, J. L., & Long, J. H. (2001). Recognition of facial affect by children and adolescents diagnosed with social phobia. *Child Psychiatry & Human Development*, 32(2), 137-145.
- Slaughter, V., Dennis, M. J., & Pritchard, M. (2002). Theory of mind and peer acceptance in preschool children. *British Journal of Developmental Psychology*, 20(4), 545–564. doi:10.1348/026151002760390945
- Smith, J. W. A., & Elley, W. B. (1997). *How children learn to read: Insights from the New Zealand experience*. Auckland, NZ: Addison Wesley Longman.
- Snyder, P. J., Nussbaum, P. D., & Robins, D. L. (2006). *Clinical neuropsychology: A pocket handbook for assessment*. American Psychological Association.
- Sporle, A., & Koea, J. (2004). Maori responsiveness in health and medical research: key issues for researchers (part 1). *The New Zealand Medical Journal (Online)*, 117(1199).
- Szczepanski, S. M., & Knight, R. T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, 83(5), 1002-1018.

- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics, 6th ed.* Boston: Pearson
- Tager-Flusberg, H., Sullivan, K., & Boshart, J. (1997). Executive functions and performance on false belief tasks. *Developmental Neuropsychology, 13*(4), 487-493.
- Tan, C. S. (2007). Test Review: Reynolds, C. R., & Kamphaus, R. W. (2004). Behavior assessment system for children (2nd ed.). Circle Pines, MN: American Guidance Service. *Assessment for Effective Intervention, 32*(2), 121–124.
<https://doi.org/10.1177/15345084070320020301>
- Tonks, J., Yates, P. J., Williams, H. W., Frampton, I., & Slater, A. (2014). Measurement Issues: Neuropsychological assessment with children and adolescents; unlocking the mysticism, methods and measures with the help of Tom Swift. *Child and Adolescent Mental Health, 19*(2), 151–158.
[doi:10.1111/camh.12043](https://doi.org/10.1111/camh.12043)
- Toplak, M. E., Bucciarelli, S. M., Jain, U., & Tannock, R. (2009). Executive functions: performance-based measures and the behavior rating inventory of executive function (BRIEF) in adolescents with attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence, 15*(1), 53–72.
<https://doi.org/10.1080/09297040802070929>
- Verger, K., Junqué, C., Jurado, M. A., Tresserras, P., Bartumeus, F., Nogues, P., & Poch, J. M. (2000). Age effects on long-term neuropsychological outcome in paediatric traumatic brain injury. *Brain Injury, 14*(6), 495-503.
- Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E. V., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling

the role of executive functions. *Developmental Neuropsychology*, 38(2), 114–25. doi:10.1080/87565641.2012.733786

- Vriezen, E. R., & Pigott, S. E. (2002). The relationship between parental report on the BRIEF and performance-based measures of executive function in children with moderate to severe traumatic brain injury. *Child Neuropsychology*, 8(4), 296–303.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development : The Truth about False Belief. *Child Development*, 72(3), 655–684.
- Wagner, S., Müller, C., Helmreich, I., Huss, M., & Tadić, A. (2015). A meta-analysis of cognitive functions in children and adolescents with major depressive disorder. *European child & adolescent psychiatry*, 24(1), 5-19.
- Wang, Z., Devine, R. T., Wong, K. K., & Hughes, C. (2016). Theory of mind and executive function during middle childhood across cultures. *Journal of experimental child psychology*, 149, 6-22.
- Wasserman, T., & Wasserman, L. D. (2013). Toward an integrated model of executive functioning in children. *Applied Neuropsychology: Child*, 2(2), 88-96.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of Theory-of-Mind development : The truth about false belief. *Child Development*, 72(3), 655–684.
- Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological science*, 17(12), 1075-1081.
- Weyandt, L. L., Willis, W. G., Swentosky, A., Wilson, K., Janusis, G. M., Chung, H. J., ... & Marshall, S. (2014). A review of the use of executive function tasks in

externalizing and internalizing disorders. In S. Goldstein, & J. A. Naglieri. (Eds.), *Handbook of executive functioning* (pp. 69-87). New York: Springer.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103-128.

Wölwer, W., Frommann, N., Halfmann, S., Piaszek, A., Streit, M., & Gaebel, W. (2005). Remediation of impairments in facial affect recognition in schizophrenia: Efficacy and specificity of a new training program. *Schizophrenia Research*, *80*(2-3), 295-303.
<https://doi.org/10.1016/j.schres.2005.07.018>

Wong, N., Beidel, D. C., Sarver, D. E., & Sims, V. (2012). Facial emotion recognition in children with high functioning autism and children with social phobia. *Child Psychiatry and Human Development*, *43*(5), 775-794.
<https://doi.org/10.1007/s10578-012-0296-z>

Yang, C., Zhang, T., Li, Z., Heeramun-Aubeeluck, A., Liu, N., Huang, N., ... Lu, Z. (2015). The relationship between facial emotion recognition and executive functions in first-episode patients with schizophrenia and their siblings. *BMC Psychiatry*, *15*(1), 241. <https://doi.org/10.1186/s12888-015-0618-3>

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: development and plasticity. *Child Development Perspectives*, *6*(4), <https://doi.org/10.1111/j.1750-8606.2012.00246.x>

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development.

APPENDIX A Information and Consent Forms



Pesticide exposure and brain function in children

Parent Information Sheet

What is this study about?

We are inviting you and your child to take part in a study looking at how much children are exposed to pesticides and whether this exposure has any effects on their brain development.

There is some concern that exposure to pesticides may affect brain function and behaviour. These effects have been found in exposed farmers and farm workers, but children may also be at risk due to the use of pesticides in the home, living close to pesticide-treated farming areas, the occupation of the parents, and to a lesser extent, through the food supply and drinking water.

This study aims to assess the level of exposure to pesticides in New Zealand children and whether this exposure has any effects on the brain and mental functions, and behaviours. To do this, we are inviting 900 children aged 5-11 years to take part in this study.

What would participation in this study involve for me?

Phase I: All participants

Your participation will involve completing three postal questionnaires on behalf of your child. The first questionnaire will ask questions about your work history, your use of pesticides, your child's diet, and your family's health. The other two questionnaires will ask questions about your child's behaviour. All three questionnaires should take no more than about 1 hour to complete. If you are interested in taking part in the study, please complete the enclosed reply form and return it to us in the freepost envelope provided. If you agree to participate, we may also ask your child's teacher to complete the two behavioural questionnaires so that we have information about your child's behaviour at school. With your permission, we will also ask your child's school to provide us with some measures of your child's academic achievements including whether they are meeting the National Standards for reading, writing, and mathematics.

If you are *not* interested in taking part, please return the reply form indicating your refusal.

For selected participants:

Phase II: Selected participants

Centre for Public Health Research

Massey University –
Wellington Campus
PO Box 756
Wellington, 6140
New Zealand

Administration

+64 (0)4 380 0602

Direct Fax

+64 (0)4 802 7120

Courier Address

Block 3, Level D

Entrance B, Wallace Street

Wellington, 6021

New Zealand

Internet

publichealth.massey.ac.nz

Research Coordinator

Jean Feary Mckenzie

+64 (0)4 979 3109
0800 990 053

j.fearymckenzie@massey.ac.nz

Later we will contact half (randomly chosen) of the children/parents again and ask the children to undergo neuropsychological tests of memory, attention and reaction time, as well as their language skills and visual processing skills. Tests will normally take place at school (but they can be done in your home if you prefer) and will be conducted by a trained researcher. The tests can be completed in less than an hour. Our research nurse will also ask your child to collect urine samples during and outside the spraying season and will ask you to collect a house dust sample. The collection of these samples will be used to examine pesticide exposure and will be fully explained by our research nurse.

What will happen with my personal information?

We will treat all of the information from the questionnaires and the information from the school about your child as strictly confidential. Each questionnaire and the results of the other tests will be entered into a database using ID numbers. The questionnaires and tests will be seen by named researchers only and when the study is completed all questionnaires will be locked away in filing cabinets which will be the responsibility of the Director of the Centre for Public Health Research. When the study has been completed, we will analyse the information, *e.g.* compare the prevalence of the effects on the brain and mental functions and behaviours in those who are exposed to pesticides compared with those who are not exposed. At the end of the study the urine and house dust samples may be tested for different chemicals related to this research. The results of the study will be published in scientific journals and a summary of the results will be provided to all study participants that have requested it. **No individual information or names will be published.**

This project has been reviewed and approved by the Central Health and Disability Ethics Committee (application ref 13/CEN/134). If you have any concerns about the conduct of this research, please contact 0800 4 ETHICS (438 442).

You have the right to:

- decline to participate
- decline to answer any of the questions
- withdraw from the study or parts of the study at any time
- be given access to a summary of the study findings when it is completed

Please contact us at the Centre for Public Health Research to discuss any queries or concerns about the study.

Thank you very much for your time in considering this study.

Study team:

- Professor Jeroen Douwes (Director), Dr Andrea 't Mannetje (Senior Research Fellow), Dr David McLean (Senior Research Fellow), Dr Amanda Eng (Research Fellow), and Jean Feary McKenzie (Research Assistant), Centre for Public Health Research, Massey University, Wellington.
- Professor Janet Leathem, Professor of Neuropsychology, School of Psychology, Massey University, Wellington.
- Kathryn McLennan, Doctoral Student, School of Psychology, Massey University, Wellington.
- Professor Neil Pearce, Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, London, UK
- Professor Jordi Sunyer, Joint Scientific Director and researcher, Centre for Research in Environmental Epidemiology, Barcelona, Spain.
- Professor Jochen Mueller, ARC Future Fellow, National Research Centre for Environmental Toxicology, University of Queensland, Coopers Plains, Australia.
- Professor Brenda Eskenazi, Director, Center for Environmental Research and Children's Health (CERCH), University of California, Berkeley, USA.



Pesticide exposure and brain function in children

Parent Reply Form

On behalf of my child:..... [Full name of child]

Name of child’s school: Year:.....

We wish to participate in the study OR

We do not wish to participate in the study (please skip to the section at the end of this form)

- We have read the Information Sheet and understand that we may ask questions at any time.
- We understand that we have the right to withdraw from the study at any time and to decline to answer any particular questions.
- We understand that no individual information or names will be published.
- We understand that the information provided will be used only for this research and publications arising from this research project.
- We understand that at the end of the project the samples may be tested for different chemicals related to this research.

We would like to be sent a summary of the study results and our child’s individual results: **yes** **no**

We consent to the investigators contacting our child’s GP if any abnormal results are found: **yes** **no**

Name of parent/guardian:

Signed (parent/guardian):

Signed (child, optional):

Date:

Postal address:

Contact phone number: (.....)

Name and address of General Practitioner:

If you are not interested in taking part, it is helpful for us to know the reason: (please tick a box)

I am not interested I don’t have time I feel that my English is not good enough to be able to complete the questionnaire

Other, please specify _____

Centre for Public Health Research

Massey University – Wellington Campus
PO Box 756
Wellington, 6140
New Zealand

Administration

+64 (0)4 380 0602

Direct Fax

+64 (0)4 802 7120

Courier Address

Block 3, Level D

Entrance B, Wallace Street

Wellington, 6021

New Zealand

Internet

publichealth.massey.ac.nz

Research Coordinator

Jean Feary Mckenzie

+64 (0)4 979 3109
0800 990 053

j.fearymckenzie@massey.ac.nz

Pesticide exposure and brain function in children

Teacher Information Sheet

What is this study about?

We have invited children at your school to take part in a study looking at how much children are exposed to pesticides and whether this exposure has any effects on their brain development.

There is some concern that exposure to pesticides may affect brain function and behaviour. These effects have been found in exposed farmers and farm workers, but children may also be at risk due to the use of pesticides in the home, living close to pesticide-treated farming areas, the occupation of the parents, and to a lesser extent, through the food supply and drinking water.

This study aims to assess the level of exposure to pesticides in New Zealand children and whether this exposure has any effects on the brain and mental functions, and behaviours. To do this, we are inviting 900 children aged 5-11 years to take part in this study.

What would participation in this study involve?

Phase I: All participants

We have sent a letter, information sheet, and consent form home with each child to their parents. We would really appreciate it if you could encourage your class to remind their parents to return the form to us indicating whether they would like to take part or not. For a small number of participants, we would like to ask you to complete two brief confidential questionnaires about the child's behaviour at school. The parents of the child have already completed a questionnaire about potential sources of pesticide exposure as well as questionnaires about the child's behaviour at home; however, it is important for the study to also collect information about the child's behaviour at school. The questionnaires that we will ask you to complete should take 15-20 minutes per child. If you are willing to take part, please complete and return the attached consent form. With parental consent, we will also ask your school to provide us with some measures of the child's academic achievements including whether they are meeting the National Standards for reading, writing, and mathematics.

For selected participants:

Phase II: Selected participants (not teachers)

Later we will contact half (randomly chosen) of the children/parents again and ask permission for their child to undergo neuropsychological tests of memory, attention and reaction time, as well as their language skills and visual processing skills. Tests will be conducted by a trained researcher and will normally take place at school. This phase of the study will not require any teacher involvement beyond releasing the children from class. At this time, our research nurse will ask the child to collect urine samples and their parents to collect a house dust sample. The

Centre for Public Health Research

Massey University –
Wellington Campus
PO Box 756
Wellington, 6140
New Zealand

Administration

+64 (0)4 380 0602

Direct Fax

+64 (0)4 380 0600

Courier Address

Block 3, Level D

Entrance B, Wallace Street

Wellington, 6021

New Zealand

Internet

publichealth.massey.ac.nz

Dr Amanda Eng

Telephone

+64 4 801 0801

0800 990 053

Email

a.j.eng@massey.ac.nz

collection of these samples will be used to examine pesticide exposure and will be fully explained by our research nurse to the children and their parents.

What will happen with the personal information?

All of the information collected from the questionnaires that you are asked to complete is strictly confidential. Each questionnaire and the results of the other tests will be entered into a database using ID numbers. The questionnaires and tests will be seen by named researchers only and when the study is completed all questionnaires will be locked away in filing cabinets which will be the responsibility of the Director of the Centre for Public Health Research. When the study has been completed, we will analyse the information, *e.g.* compare the prevalence of the effects on the brain and mental functions and behaviours in those who are exposed to pesticides compared with those who are not exposed. At the end of the study the urine and house dust samples may be tested for different chemicals related to this research. The results of the study will be published in scientific journals and a summary of the results will be provided to all study participants that have requested it. **No individual information or names will be published.**

This project has been reviewed and approved by the Central Health and Disability Ethics Committee (application ref 13/CEN/134). If you have any concerns about the conduct of this research, please contact 0800 4 ETHICS (438 442).

You have the right to:

- decline to participate
- decline to answer any of the questions
- withdraw from the study or parts of the study at any time
- be given access to a summary of the study findings when it is completed

Please contact us at the Centre for Public Health Research to discuss any queries or concerns about the study.

Thank you very much for your time in considering this study.

Study team:

- Professor Jeroen Douwes (Director), Dr Andrea 't Mannetje (Senior Research Fellow), Dr David McLean (Senior Research Fellow), Jean Feary Mckenzie (Research Assistant) and Dr Amanda Eng (Research Fellow), Centre for Public Health Research, Massey University, Wellington.
- Professor Janet Leathem, Professor of Neuropsychology, School of Psychology, Massey University, Wellington.
- Kate Ross-McAlpine & Yanis Brinkmann, Doctoral Students, School of Psychology, Massey University, Wellington.
- Professor Neil Pearce, Professor of Epidemiology and Biostatistics, London School of Hygiene and Tropical Medicine, London, UK
- Professor Jordi Sunyer, Joint Scientific Director and researcher, Centre for Research in Environmental Epidemiology, Barcelona, Spain.
- Professor Jochen Mueller, ARC Future Fellow, National Research Centre for Environmental Toxicology, University of Queensland, Coopers Plains, Australia.
- Professor Brenda Eskenazi, Director, Center for Environmental Research and Children's Health (CERCH), University of California, Berkeley, USA.



Pesticide exposure and brain function in children

Teacher Consent Form

- I wish / do not wish to participate in the study (please delete one).
- I have read the Information Sheet and understand that I may ask questions at any time.
- I understand that I have the right to withdraw from the study at any time and to decline to answer any particular questions.
- I understand that no individual information or names will be published.
- I understand that the information provided will be used only for this research and publications arising from this research project.

Centre for Public Health Research

Massey University –
Wellington Campus
PO Box 756
Wellington, 6140
New Zealand

Administration

+64 (0)4 380 0602

Direct Fax

+64 (0)4 380 0600

Courier Address

Block 3, Level D

Entrance B, Wallace Street

Wellington, 6021

New Zealand

Internet

publichealth.massey.ac.nz

Dr Amanda Eng

Telephone

+64 4 801 0801

0800 990 053

Email

a.j.eng@massey.ac.nz

I would like to be sent a summary of the study results: yes no

Name:.....

Signed:

Name of school:

Class year:.....

Date:

APPENDIX B

Adaptions to Selected Measures

Attachment to BASC-2 and BRIEF questionnaires

The following statement was attached as a sticker to the front of all BASC-2 and BRIEF questionnaires:

These questionnaires were developed in the United States and are used worldwide. As such, some of the language and spelling may differ from NZ English. Where the word “grade” is used, please substitute “mark”. Where the word “recess” is used, please substitute “morning tea”. Where the word “closet” is used, please substitute “wardrobe”.

Table B1

List of cultural adaptations made to the NEPSY-II ToM task

Original	Substitution
André	Andrew
Mom	Mum
Ming	Mandy
Sheryl	Sophie
Luz	Liz
Brandon	Matthew
Mom	Mum
Reggie	Reagan
Audrey	Ashleigh
Fun House	Haunted House
Eric	Aaron
Laurie Lamb	Laura Lamb
Recess	Morning Tea
Denise	Daisy
Mama	Their Mum
Laurie Lamb	Laura Lamb
Uncle Carlos	Uncle Charlie

APPENDIX C
Non-Parametric Statistics

Table C1

Non-parametric analyses examining whether NEPSY-II scores differ as a function of examiner

	Independent Samples Kruskal-Wallis Test
AR	.77
ToM	.00*
Animal Sorting	.35
Naming	.49
Inhibition	.77
Switching	.22

*Significant difference

Table C2

Non-parametric analyses examining whether BASC-2 and BRIEF scores differ from US norms

	One-Sample Wilcoxon Signed Rank Test	
BASC-2	Parent	Teacher
Externalising Probs	.00*	.00*
Internalising Probs	.00*	.08
BSI	.00*	.00*
Adaptive Skills	.00*	.00*
BRIEF		
BRI	.00*	.00*
MI	.00*	.02*
GEC	.00*	.01*

*Significant difference

Table C3

Non-parametric analyses examining whether NEPSY-II scores differ from US norms

	One-Sample Wilcoxon Signed Rank Test
AR	.94
ToM	.74
Animal Sorting	.00*
Naming	.00*
Inhibition	.04*
Switching	.00*

*Significant difference

Table C4

Non-parametric analyses examining whether BASC-2 and BRIEF scores differ across categories of decile

BASC-2	Parent	Teacher
Externalising Probs	.04*	.35
Internalising Probs	.17	.52
BSI	.21	.12
Adaptive Skills	.06	.04*
BRIEF		
BRI	.08	.24
MI	.00*	.09
GEC	.00*	.06

Note: Decile categories determined by ERO (2015).

*Significant difference

Table C5

Non-parametric analyses examining whether BASC-2 and BRIEF scores differ as a function of gender

BASC-2	Parent	Teacher
Externalising Probs	.98	.07
Internalising Probs	.97	.18
BSI	.84	.62
Adaptive Skills	.07	.17
<hr/>		
BRIEF		
<hr/>		
BRI	.80	.55
MI	.83	.41
GEC	.82	.46

*Significant difference

APPENDIX D

Reflections on Research Process and its Application to Clinical Practice

A case study presented in partial fulfilment of

The degree of

Doctorate of Clinical Psychology

Kathryn McLennan

2016

This case study represents the work of Kathryn McLennan during her doctoral studies. The research component, supervised by Prof. Janet Leathem, was conducted in coordination with the Centre of Public Health Research. Special thanks go to the participants in the study, their parents, teachers and schools for their contribution to the research. Practical experience was obtained during my internship at Talking Point in New Plymouth in 2016, and prior placements at the Massey University Psychology Clinic and Mental Health Admissions Unit at Nelson Hospital. Any client information included in the case study has been changed to protect their anonymity.

Candidate: Kathryn McLennan.....Date:.....

Supervisor: Janet Leathem.....Date:.....

Abstract

The Doctor of Clinical Psychology Programme requires that training clinicians complete research in the form of a doctoral level thesis (attached) as well as coursework and clinical practice. Coursework components include a case study completed during the trainee's internship year that describes how research knowledge has been used to inform practice. This case study is designed to meet that criteria. It is presented as an adjunct to the thesis, and is written with the premise that the reader has prior knowledge of the research completed. Outlined in the case study are the author's reflections on the research process, experiences working with children, with a team and with schools, and the many ways in which completing this research has informed clinical practice.

Introduction

Clinical Psychologists should be able to effectively combine research and practice (Jones & Mehr, 2007). In our training we are taught to use the scientist-practitioner model, which posits that professional psychologists should be able to not only consume, evaluate and apply the best available scientific research to their clinical practice but also contribute to that knowledge base in some way to move the field forward (Jones & Mehr, 2007). The attached thesis represents just one effort to contribute to the existing knowledge base, providing new ideas about the way we should interpret multi-method findings from executive and social functioning assessments with NZ school-aged children. This case study however is designed to outline my reflections on the research process and how this has impacted on my clinical practice.

As noted in the body of the thesis, this research was completed in the context of a larger study run by the Centre of Public Health Research (CPHR) that aimed to evaluate the effects of pesticide exposure on the neuropsychological functioning of children. Working and collaborating with the CPHR research team provided some insight into working within a multi-disciplinary team, as is common in clinical practice. My role in the team was facilitating initial engagement with the schools and the administration and scoring of neuropsychological measures. Neuropsychological assessments involved the administration of cognitive tests completed with participating children (aged 5-12) and the completion of pen and paper questionnaires by their parents and teachers. As such, working on this project provided the opportunity to work with children as well as a range of other individuals involved in their development. I was also able to expand on previous research experience working within the schooling system and develop my clinical skills in administering, scoring, and interpreting

psychometric tests. The combination of developing research skills, the knowledge gained from my research project and my experiences of working with children, with schools, and within a multi-disciplinary team, have all contributed to my development as a clinician.

The first section of this study will discuss specific research skills that I have been able to apply to my clinical practice. This is then followed by a discussion of how I was able to apply knowledge gained from my thesis research, specifically regarding neuropsychological assessment and cognitive development in childhood. Final reflections will focus on my experiences of working with children, working within a school environment, and working with a team, and how these experiences have shaped my clinical practice.

Research skills

Lane and Corrie (2007) proposed a four-theme framework outlining the skills involved in being a scientist-practitioner and applying research knowledge to practice. The first theme proposed is the development and use of critical thinking skills (judgment, reasoning, problem-solving and decision-making), the ability to understand the implications of using theoretical concepts, and the ability to critically evaluate how research findings and methodology might translate to the current client situation. Over the course of my doctoral studies I have been taught to reason and justify decision-making. In research this began by critically evaluating articles and their strengths and weaknesses, identifying an area where there was a need for further study, and being able to justify why that research was needed.

In practice this was a vital skill as I had to be aware of and understand the evidence base for different clinical issues, different assessment tools, and different interventions, as well as being able to evaluate and rationalise the extent to which they might be appropriate for the client (e.g., cultural differences from standardization sample; co-morbidity). Being able to critically evaluate research was one way to ensure clients were receiving the best possible evidence-based treatment.

The second theme proposed is being able to draw together assessment information and up to date empirical findings to inform procedure. In research, this skill was used in summarizing a review of the literature and its gaps, and being able to identify the desired aims of the proposed study and how these would be achieved. In practice, these skills were necessary to be able to develop a formulation for my clients; a summary of their experience and assessment information (results from psychometric assessment, client presentation etc), grounded in psychological theory and knowledge, that was meaningful for the client and informed treatment.

The third theme proposed is the ability to apply that knowledge effectively. In research this meant developing a method that adequately answered the research questions. In practice, this meant developing the skills to apply appropriate therapeutic interventions. Finally, the fourth component described is the ability to monitor progress, to evaluate, reflect and respond appropriately to feedback, adapting where necessary. Through my research and clinical practice, I have learnt what an important role this plays in being a good practitioner. These skills require the researcher/practitioner to reflect on what is working well and areas where there is a need for improvement or change.

In research, this was evident in the evolution of my thesis. The initial concept for this thesis had involved developing a model of the relationship between EF and ToM. However, on completion of the literature review and the results of the previous studies it became apparent that it no longer made sense to do this. The thesis aims were evaluated and adapted in response to this new evidence and the rest of the thesis changed accordingly.

With clients, being able to scientifically monitor progress ensures that the client is actually progressing. If not, these checks signal that it is time to re-evaluate and adapt the treatment plan. These skills though, are important for all areas of clinical practice. Having the ability to identify areas of personal weakness and being able to take on board feedback opens up opportunities to improve the way you work professionally and ultimately improves the client's experience. These skills are also why supervision and professional development are so important.

In thinking about this topic, I also reflected on problems with confirmatory hypothesis testing (e.g., Spengler, Strohmer, Dixon & Shivy, 1995) and how important it is in practice not to just assume that our initial hypothesis is always correct. Having completed my internship year, it has become more apparent to me how important it is to view the formulation as a working hypothesis as this, like my thesis, is often updated and adapted as more information is learnt about the client.

The nature of clinical work means that there will always be areas in which I am less familiar and all of these skills will therefore be important throughout my career. While my experiences with research and with internship and placements have helped me to hone these skills, I am aware that practice and development of these skills will be ongoing.

Knowledge Gained Through Thesis Research

Neuropsychological assessment. Prior to beginning my studies at Massey University, I had limited foundational knowledge on neuropsychological assessment. I became interested in the topic during a fourth year neuropsychology paper and through discussions with my now-supervisor on possible research options. The wealth of literature on child and adult neurocognitive functioning I had to consume for the purposes of completing my thesis provided a strong theoretical knowledge that I was able to apply when completing neuropsychological assessments in my internship year.

As an example, in one neuropsychological assessment, I was faced with a client who sustained a severe TBI many years ago. His behaviour would suggest that there were underlying executive deficits, as he presented with difficulties of inhibition and initiation, he was perseverative in conversation, and there was a documented personality change following his TBI which had had a profound impact on his social and emotional functioning. A behavior rating scale indicated that the client had executive deficits. Despite these difficulties, this client performed well on objective tests of executive function. Before conducting research in this area I had not realised there was such a discrepancy between client report/rating scales and performance measures. Based on information gained through my research, as well as additional assessment and research (e.g., Toplak, West & Stanovich, 2013), I was able to identify that the client only had difficulties in unstructured environments. As objective measures were structured he was able to perform at a level consistent with premorbid estimates of function. Completing research in this area has made me more thoughtful in my interpretation of discrepancies between measures.

On a more practical level, completing this research also provided a wealth of experience administering, scoring, and interpreting findings from neuropsychological tests that was invaluable preparation for internship year. Scoring of the BASC-2 and BRIEF forms took a substantial amount of time as initially we did not have computer scoring. By hand scoring forms though, I was able to gauge a greater appreciation of the individual, comparing both parent and teacher responses for a particular child. It was also by hand scoring forms that I first noticed the number of critical items being identified by respondents. Critical items had not been considered when the ethics application was first drafted, and yet clinically and ethically I felt an obligation to follow these items up. Discussing these clients with my supervisor was incredibly valuable in helping me conceptualize and critically evaluate what that profile might mean for that particular client. These conversations also gave me experience more broadly of the value of supervision. The follow up conversations I had with parents were my first real exposure to asking about risk in a clinical context, and something which prepared me well for my internship year. I recall feeling initially nervous about making these calls and about asking such direct questions about a child I may not necessarily have met. The more practiced I became with these conversations, the more comfortable I became with asking risk questions directly and professionally. I was able to gain confidence having these conversations which are so crucial in clinical work.

Cognitive development. Prior to conducting this research, I had only thought about ToM within the context of early development or as a deficit in clinical disorders such as Autism. Working in this area has expanded my knowledge of the construct and I have begun thinking about it in relation to EF performance. During my internship year I was also able to apply these skills in both assessment of social cognition and through

cognitive rehabilitation work. For example, one client I had the opportunity to work with presented with social perception deficits and I was able to work with him on developing affect recognition and social skills. This work was greatly enhanced by knowledge gained in my research.

Working with Children

The assessments that I completed as part of this research project were with typically developing children, and yet (as to be expected) there were substantial differences in performance, presentation, ability to sustain attention and levels of engagement across individuals. Even though we were administering the same tasks, ideally in the same way each time, each assessment was unique as a result of the child being assessed. Some children were quiet and focused well on tasks but with minimal spontaneous interaction, some were more interactive, and still others were highly distractible. Having these experiences was good preparation for conducting neuropsychological assessments with clients as I had to learn how to work with children in different stages of development and with differing levels of engagement. It made me more aware that clients I work with in clinical practice may present with more pronounced difficulties, and made me think about the range of internal and external factors that can impact on neuropsychological test performance and how I would work differently with clients as a function of those different factors.

One aspect of working with children that I found challenging initially was in inhibiting my own automatic response to give praise and tell the child they were doing well. This was especially difficult when children would ask if they had correctly answered the question. I found that it took a lot of work and practice to deliver more

standardized responses. In my internship I found that adults too would ask for feedback on their performance, so learning how to respond to these situations in my research was directly transferable to clinical work.

Working with schools

Having completed research in schools before, and coming from a family of teachers, I felt I was able to bring to the team a greater understanding of the demands placed on teachers and how important it was that our expectations for the level of work they were being asked to complete were realistic. Developing a good relationship with the school from the beginning was an integral part of ensuring a good response rate, particularly as it was up to Principals to promote the study within their school. It was clear that when Principals were interested and engaged in the research, the whole process went more smoothly.

Schools are a dynamic environment, and working within the school environment required a flexible, adaptable approach, not unlike clinical work. The more time I spent doing this work, I realized how important it was to call ahead and confirm that the child(ren) you were hoping to see were actually there. Particularly in the beginning, we had several incidents where we had driven half an hour to arrive at the school with only one or two children left to see and one of them was sick, away at swimming, or attending an important assembly and was not available to complete the assessment. Frequently we would go to collect somebody from class and find they were in a different reading group on the other side of the school. These experiences led me to reflect that as a Psychologist, when completing school observations with children, these

little organizational tips will be even more important and one needs to take into account the changing environment.

It always impressed me when visiting schools how accommodating teachers were of interruptions. My experience of working with teachers and within the schools made me reflect on how important it is to be aware of the classroom environment that both teacher and child are working in and how this impacts on the child's learning as well as the teacher and how they manage or interpret behavior. It also caused me to consider how important school observations would be in clinical practice, seeing how the child interacts with peers, how they behave in class, and the way in which the teacher responds to that behavior, as this is information that may otherwise be missed in a clinical assessment.

Working in a team (multi-disciplinary)

One of the benefits of working in a team is that everybody brings their own background knowledge and unique skills. The current research was multidisciplinary in that it brought together researchers from Psychology and Public Health. The Public Health team had extensive experience of working on large scale research models as well as knowledge of pesticide exposure and its related effects. Researchers from Psychology brought experience and knowledge of neuropsychological assessment and cognitive functioning in children. This kind of research would not have been able to be completed had these two teams not come together in collaboration.

One of my early observations in working with a different team was the different approaches taken with respect to research. As an example, in the early stages of the project I was involved in the review process of the information and consent forms and

the Pesticide Questionnaire that parents would be asked to complete. The Public Health Team had completed large-scale exposure studies in the past and had compiled an extensive questionnaire to gather information regarding child and parent background and the various genetic, developmental and environmental factors that may contribute to findings of a relationship between pesticide exposure and neuropsychological functioning. On reviewing the questionnaire however, I couldn't help but think of the parents who would be completing it, as the measure, while exhaustive, would take approximately an hour to complete. I wondered whether a busy parent with no great incentive to complete the questionnaire would simply select out of the study. In discussing this with the team we were able to revise and reduce the workload for parents involved in the study. In a similar way, myself and others from Psychology were able to contribute knowledge of child development and language when it came to developing child information and consent forms appropriate for the developmental level of children we were recruiting.

One of the disadvantages I found of working in a large team was that there was less oversight when it came to entering data meaning there was a higher probability of human error. One such instance was when it was identified that there had been errors in scoring and entering, so on receiving this information all data were re-scored and re-entered. What ended up happening however, was that the "errors" that had been noted were the result of another person using a different set of norms. As a result, several weeks were spent unnecessarily re-scoring and re-entering data that had actually been correct in the first place. Thankfully an original data set had been kept and this was able to be used for analysis. This experience reinforced for me the need to have checks in place for scoring and entering of data and very specific protocols that everyone can follow. In future I would be more careful to ensure data were being scored correctly,

and to check that errors were actual errors before re-scoring. It also made me more aware that processes have to be transparent for somebody else to follow.

Working in a larger team also meant that organisation was more difficult and aspects of completing research, particularly in the beginning, took much longer to complete than expected as all involved in the study had to sign off on the project. Given that study collaborators for the research project were placed all over the globe, this proved challenging for me, knowing that my research was time-limited. This experience though has helped me be more adaptable, flexible and patient as these obstacles were out of my control. It also made me think more about how I might manage these types of obstacles when working with others involved in a client's care. For instance, if my work with a client requires me to gather information from another agency, how will I manage situations in which their priorities are different? Working with this team has also given me experience with shared decision-making and the practicalities of working with a multi-disciplinary team. These experiences have already proved useful in preparing me for clinical work with a mental health team.

Other Learning

During the research development stage, the BASC-2, BRIEF, and cognitive tasks were selected for use with the larger study. Due to the wealth of research on these measures in recent years, several updates occurred over the course of the research period that were unavoidable. First, the BASC-3 and BRIEF-2 were developed and released and second, computer scoring was made available that substantially reduced the amount of time spent scoring assessments. Had I known in the beginning that these changes were underway, I may have prioritized my time differently so that data

collection/scoring could be completed in a shorter timeframe later in the process. As above however, I have had to accept that these things happen in research and practice. These experiences have helped me become more relaxed about change, and as a result I have been able to be more adaptable and work more effectively as a practitioner.

Discussion

The premise of this case study was to reflect on the process of conducting research, my experiences of working with a team, with child participants and with schools more generally, and finally to convey how those experiences contributed to my practice as a clinician. As discussed there has been several ways in which I have been able to use the scientist-practitioner approach to apply research skills to practice. I have learnt to be inquisitive and to think critically, skills that have helped me become a better practitioner. During my training I have also been fortunate to have received other opportunities in which to disseminate research ideas, presenting on both my own research and a range of other topics. Most importantly, I have learnt the value of being a reflective practitioner. The skills and knowledge I have gained from completing this research will continue to inform my practice and growth as a Clinical Psychologist.

References

- Jones, J. L., & Mehr, S. L. (2007). Foundations and assumptions of the scientist practitioner model. *American Behavioral Scientist, 50*(6), 766-771.
- Lane, D. A., & Corrie, S. (2007). *The modern scientist-practitioner: A guide to practice in psychology*. Routledge.
- Spengler, P. M., Strohmer, D. C., Dixon, D. N., & Shivy, V. A. (1995). A scientist practitioner model of psychological assessment: Implications for training, practice and research. *The Counseling Psychologist, 23*(3), 506-534.
- Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry, 54*(2), 131-143.