

Developing and Piloting a New Measure of Executive Functioning for
Children with Autism Spectrum Disorder (ASD)

Eleonore Bristow

D.Clin.Psy thesis (Volume 1), 2016

University College London

UCL Doctorate in Clinical Psychology

Thesis declaration form

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

Signature:

Name:

Date:

Overview

This thesis focuses on executive functioning in children with Autism Spectrum Disorder (ASD). The literature review (Part 1) uses meta-analytic techniques to assess whether planning impairments are present in children with ASD. It also has the secondary aim of exploring the factors that may be associated with planning task performance. Specifically, the impact of participant age, study quality and computer vs. researcher administration of tasks are investigated.

Part 2, the empirical research paper, describes the development and subsequent pilot of a new measure: the Ecologically-Valid Test of Executive Dysfunction (Eco-TED) for children with ASD. This project was conducted jointly with another trainee and therefore only three of the seven final Eco-TED subtests are given consideration here: the Consent Form Test and the School Bag and Lego Tasks. Part 2 examines the tests' ability to detect executive difficulties in children with ASD, as well as their test-retest reliability and ecological and criterion validity.

The critical appraisal (Part 3) sets out a number of reflections on the process of carrying out the literature review and the empirical research study. In particular, the nature of ASD and the impact this has on how research is conducted is discussed. In addition, the tasks of neuropsychological test development and the supervision of a small team of research assistants are considered. Finally, recommendations for future development of the Eco-TED are put forward.

Contents

Overview.....	3
List of Tables.....	9
List of Figures.....	10
Acknowledgements.....	11
Part 1: Literature Review.....	12
1.1 Abstract.....	13
1.2 Introduction.....	14
1.2.1 Autism Spectrum Disorder (ASD).....	14
1.2.2 Executive functioning in ASD.....	14
1.2.3 Planning.....	15
1.2.4 Measures of planning.....	16
1.2.5 Planning in ASD.....	16
1.2.6 Research aims.....	18
1.3 Method.....	19
1.3.1 Data sources and study inclusion.....	19
1.3.2 Study quality.....	20
1.3.3 Planning measures.....	21
1.3.4 Effect size calculation.....	22
1.3.5 Statistical procedures.....	23
1.4 Results.....	26
1.4.1. Corpus of studies.....	26
1.4.2 Study quality.....	27
1.4.3 Meta-analysis results.....	27

1.4.4 Publication bias.....	27
1.4.5 Factors affecting planning ability.....	32
1.4.5.1 Mode of administration.....	32
1.4.5.2 Participant age.....	32
1.4.5.3 Study quality.....	33
1.5 Discussion.....	33
1.6 References.....	40
Part 2: Empirical Paper.....	47
2.1 Abstract.....	48
2.2 Introduction.....	49
2.3 Method.....	54
2.3.1 Participants.....	54
2.3.2 Procedure.....	56
2.3.3 Measures.....	57
2.3.3.1 Eco-TED.....	57
2.3.3.1.1 <i>Consent Form Test</i>	59
2.3.3.1.2 <i>School Bag Task</i>	60
2.3.3.1.3 <i>Lego Task</i>	61
2.3.3.2 Neuropsychological measures.....	63
2.3.3.3 Parent-report questionnaires.....	64
2.3.3.3.1 <i>The Behaviour Rating Inventory of Executive Functions (BRIEF)</i>	64
2.3.3.3.2 <i>Social and Communication Disorders Checklist (SCDC)</i>	64

2.3.3.3.3	<i>Strengths and Difficulties Questionnaire (SDQ)</i>	64
2.3.4.	Data analysis.....	65
2.3.4.1	Data analysis procedures.....	65
2.3.4.1.1	Group differences.....	65
2.3.4.1.1.1	Eco-TED raw scores.....	65
2.3.4.1.1.2	Eco-TED deficit scores.....	66
2.3.4.1.1.3	School Bag Task individual items.....	68
2.3.4.1.1.4	Individual variation in performance...	68
2.3.4.1.2	Correlations.....	69
2.4	Results.....	69
2.4.1.	Group differences.....	69
2.4.1.1	SDQ and SCDC.....	69
2.4.1.2	Executive functioning.....	69
2.4.1.2.1	BRIEF.....	69
2.4.1.2.2	BADS-C.....	69
2.4.1.2.3	Eco-TED.....	69
2.4.1.2.3.1	Eco-TED raw scores.....	69
2.4.1.2.3.2	Eco-TED deficit scores.....	70
2.4.1.2.3.3	School Bag Task individual items.....	70
2.4.1.2.3.4	Individual variation in performance...	70
2.4.2	Correlations.....	76
2.4.2.1	Correlations with SDQ and SCDC.....	76
2.4.2.2	Correlations with IQ and age.....	76

2.4.2.3 Correlations with established measures of executive functioning.....	76
2.4.2.4 Test-retest reliability.....	77
2.4.2.5 Post hoc correlations.....	77
2.5 Discussion.....	80
2.5.1. The Eco-TED pilot: Summary and interpretation of results....	80
2.5.1.1 Consent Form Test.....	81
2.5.1.2 School Bag Task.....	83
2.5.1.3 Lego Task.....	86
2.5.1.4 Test-retest reliability.....	88
2.5.2 Limitations and strengths of the study.....	89
2.5.2.1 Limitations.....	89
2.5.2.2 Strengths.....	90
2.5.3 Conclusion.....	91
2.6 References.....	93
Part 3: Critical Appraisal.....	101
3.1 Introduction.....	102
3.2 Carrying out research with individuals with ASD.....	102
3.2.1 Confounding variables.....	102
3.2.2 Heterogeneity among individuals with ASD.....	104
3.2.3 Methodological challenges.....	105
3.2.4 A deficit narrative.....	106
3.3 Neuropsychological test development.....	107
3.3.1 Practicalities.....	108

3.3.2 Piloting and service user involvement.....	109
3.3.3 Taking a ‘bottom-up’ approach to task development.....	109
3.3.4 Developing instruments for use in specific populations.....	111
3.3.5 Control Group Choice.....	111
3.4 Supervision of research assistants.....	113
3.4.1 Successes.....	113
3.4.2 Areas for improvement.....	114
3.5 The future of the Eco-TED.....	115
3.5.1 A larger sample.....	115
3.5.2 Item 6 of the School Bag Task.....	116
3.5.3 Questionnaire development.....	116
3.6 Conclusion.....	116
3.7 References.....	118

Appendices

Appendix I	Critical Appraisal Skills Programme checklist.....	121
Appendix II	Ethical approval document.....	128
Appendix III	Information sheets and consent forms.....	133
Appendix IV	Initial contact letters for parents.....	145
Appendix V	Eco-TED Consent Form Test, School Bag Task and Lego Tasks script.....	148
Appendix VI	Trainee contribution to the project.....	156
Appendix VII	SDQ.....	158
Appendix VIII	School Bag Task individual items analysis results.....	161

List of Tables

Part 1: Literature Review

Table 1.1 Summary of the tower task variants.....	24
Table 1.2 Summary of the dependent measures used for the purposes of the meta-analysis.....	26
Table 1.3 Summary of studies included in the meta-analysis.....	28
Table 1.4 Critical Appraisal Skills Programme (CASP, 2014) checklist scores.....	30

Part 2: Empirical Paper

Table 2.1 Mean (and standard deviations) for participant characteristics in both groups.....	57
Table 2.2 Items from the 3Di that formed the basis of the development of the Eco-TED subtests.....	59
Table 2.3 Group differences, means, standard deviations, p values and effect sizes for the SCDC, SDQ, Zoo Map and Six Part Tests, and the BRIEF.....	71
Table 2.4 Group differences, means, standard deviations, p values and effect sizes for the raw and deficit Eco-TED scores.....	72
Table 2.5 Correlation coefficients and significance values by group for the Eco-TED measures and SDQ Overall Stress, SCDC Total, Zoo Map and Six Part Tests, and BRIEF Global Executive Composite scores.....	78
Table 2.6 Correlation coefficients for the initial Eco-TED raw scores and retest Eco-TED raw scores.....	80

List of Figures

Part 1: Literature Review

- Figure 1.1 Literature search strategy.....25
- Figure 1.2 A forest plot of planning performance effect sizes and confidence intervals for each study included in the meta-analysis.....31
- Figure 1.3 A funnel plot to explore for publication bias.....31

Part 2: Empirical Paper

- Figure 2.1 Control participants' profile of performance across the 34 Eco-TED sub-measures for which z -scores could be calculated.....74
- Figure 2.2 ASD participants' profile of performance across the 34 Eco-TED sub-measures for which z -scores could be calculated.....75

Acknowledgements

I would like to thank my supervisors Dr Will Mandy and Professor Paul Burgess for their unwavering support and guidance throughout the project. I am so grateful to have had the opportunity to be supervised by two experts in this field. I have not only learnt a great deal from them both but have also enjoyed working with them enormously.

I would also like to thank the other members of the research team. Eleonora, Aurélie and Salma have worked hard to support this study and their contribution has been significant. I also want to thank Great Ormond Street Hospital for their involvement, and Rob Hickman for his enthusiasm and invaluable contribution to our clinical recruitment. Additionally, I am indebted to my friend Carrie King who was instrumental in the recruitment of our control group. Most of all, I'd like to thank Jodie for being my co-researcher and friend – I could not have done it without you and feel so lucky to have been on your team!

I am so thankful for the support I have received from family and friends throughout this process, particularly Matt who has been both a voice of calm encouragement and the perfect distraction when I've needed it. I am also so grateful to my parents for having faith in me, and my career choice, and for supporting me always in achieving my ambitions.

Finally, I want to express my gratitude to our participants and their families for taking part in the research study. I have been touched and motivated by their interest in the research, which has ranged from personal to philanthropic. It has been a unique opportunity and great privilege to meet these families.

Part 1: Literature Review

Planning in Children with Autism Spectrum Disorder (ASD): A Meta-Analysis

1. 1 Abstract

Aims: To assess whether planning impairments are present in children with ASD and explore the factors that may be associated with planning task performance.

Method: PSYCinfo, EMBASE and MEDLINE were searched using terms related to autism, executive functioning, planning, and children. The search yielded 149 papers, 27 of which were retrieved but only 17 of which were selected for the final analysis on the basis of our inclusion criteria. As tower planning tasks are analogous, the review was limited to studies using these as a way of reducing the heterogeneity of the data. Standardised effect sizes of tower task performance were calculated and the meta-analysis conducted using a random effects model. Categorical analysis was used to explore factors (participant age, study quality, computer vs. researcher administration of tasks) influencing planning ability.

Results: We found an overall significant negative effect of ASD ($d = -0.65$) on performance in tower planning tasks. We found no effect for age or study quality on task performance. Studies using computerised versions of planning tasks showed only a trend effect of ASD, while those using researcher-administered versions showed a significant effect.

Conclusion: This review found evidence for impaired performance in tower planning tasks amongst children with ASD. Our findings offer tentative support for the idea that the social demands of laboratory tasks may affect executive performance. However, it is difficult to draw definitive conclusions about planning per se as tower tasks have a number of limitations, perhaps the most concerning of which is their lack of ecological validity.

1.2 Introduction

1.2.1 Autism Spectrum Disorder (ASD)

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental condition characterised by impaired social communication, restricted interests and repetitive behaviours. The Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5; American Psychiatric Association, 2013) diagnostic criteria include social interaction difficulties; specifically, deficits in socio-emotional reciprocity, non-verbal communication, and difficulties in forming and maintaining relationships. In order to meet the diagnostic threshold, individuals must also display at least two of the following: repetitive and stereotyped behaviours and speech; excessive adherence to routines and rituals; highly fixated and restricted interests; and sensory abnormalities. ASD is now understood as a dimensional disorder and diagnosed individuals may present in a wide variety of ways in terms of their behaviour and the severity of their symptoms (Nydén, Hagberg, Goussé, & Rastam, 2011). ASD is present in at least 1% of the population and occurs in three times as many males as females (Baird et al., 2006). Although intellectual disability affects 25-40% of cases (Baird et al., 2000; Baird et al., 2006; Chakrabarti & Fombonne, 2001), the majority of individuals with ASD have intellectual abilities within the normal ranges.

1.2.2 Executive functioning in ASD

Cognitive theories of ASD, and therefore an interest in the cognitive phenotypes of individuals with ASD, have dominated psychological research in this area (Rajendran & Mitchell, 2007). One of the current prominent cognitive accounts of ASD is the theory of executive dysfunction. This theory arose due to observed similarities between patients with acquired frontal lobe damage and those with ASD (e.g. Ozonoff, Pennington, & Rogers, 1991). The model directly links

frontal and executive functioning with the symptoms of ASD and seeks to account for both the social and non-social behaviours seen in individuals with ASD in terms of executive dysfunction. For example, it understands symptoms such as rigidity and perseveration in terms of difficulties with the executive functions of initiation of action and shifting set (Hill, 2004a). One criticism of the executive functioning account of ASD is that it lacks specificity, since deficits in executive functioning are implicated in a wide range of disorders including Parkinson's disease (Taylor, Saint-Cyr, & Lang, 1986), attention deficit hyperactivity disorder (ADHD; Barkley, 1998) and schizophrenia (Weinberger & Gallhofer, 1997). It is as yet unclear which aspects of executive functioning are key in autism, or indeed if any disorder-specific pattern exists. Researchers have therefore started to move towards trying to identify the specific executive functioning profile for ASD, or at least the types of executive task that are typically failed by people with ASD (White, Burgess, & Hill, 2009).

Planning has been proposed and investigated as one of the key executive domains implicated in ASD. The literature on planning in ASD is discussed below, but planning and measures of planning are outlined first.

1.2.3 Planning

The ability to plan is an important cognitive skill and, as is the case with shifting set, inhibition, and other skills that fall under the 'executive function' umbrella, it enables us to respond in adaptive ways to complex or novel situations (Happé, Booth, Charlton, & Hughes, 2006). These executive domains allow us to control our behaviour in order to attain an overarching goal, and without them our behaviour can become disorganised and incoherent (Duncan, 1986).

Hill (2004a) describes planning as a "complex, dynamic operation in which a sequence of planned actions must be constantly monitored, re-evaluated, and

updated. This requires the conceptualisation of changes from the current situation, looking ahead by taking an objective and abstracted approach to identify alternatives, making choices, and then implementing the plan and revising it accordingly”. This account captures the numerous facets of planning and therefore the complexity of the task.

1.2.4 Measures of planning

Although a number of different measures have been used to assess planning abilities in ASD, tower tasks have been the most widely used. These tasks require participants to rearrange a number of disks, balls, or rings into a given target configuration whilst following a particular set of rules (e.g. move only one ring at a time). A number of tower variants appear in the literature: the Tower of Hanoi; the Tower of London; the Stockings of Cambridge subtest of the Cambridge Neuropsychological Test Automated Battery (CANTAB); the Delis–Kaplan Executive Functioning System Tower (D-KEFS; Delis, Kaplan, & Kramer, 2001); and the NEPSY Tower (Korkman, Kirk, & Kemp, 1998).

1.2.5 Planning in ASD

In a qualitative, narrative review focusing on planning, inhibition, mental flexibility, generativity and action monitoring, Hill (2004a) argued for the existence of significant planning deficits in individuals with ASD, supporting earlier findings by Sergeant, Geurts, & Oosterlaan (2002) and Pennington & Ozonoff (1996). More specifically, Hill (2004a) reported that, overall, children and adults “exhibit a planning impairment, as assessed by Tower tasks” (p. 196). Hill's (2004a) description of an “overall” effect was based on a qualitative synthesis of the literature and it is important to note that a number of researchers have found negative results using tower tasks to measure planning ability in ASD (e.g. Corbett,

Constantine, Hendren, Roche, & Ozonoff, 2009; Goldberg et al., 2005; Happé, Booth, Charlton, & Hughes, 2006; Sinzig, Morsch, Bruning, Schmidt, & Lehmkuhl, 2008). In addition, considering the tower task literature alongside evidence from studies using alternative planning measures muddies the water further. Planning impairments at the level of motor control have been found among ASD participants using the Milner mazes task (Milner, 1965; Prior & Hoffmann, 1990) and Luria's bar task (Hughes, 1996), but not when assessed with the mazes task from the Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1989; Pellicano, 2007). There are therefore a number of discrepancies in the literature and it remains unclear whether individuals with ASD are typically impaired with regard to planning.

Some of the inconsistencies in the literature may simply be accounted for by the variety of measures that have been used to assess planning in ASD in the sense that different tasks may assess different aspects of planning. However, as mentioned above, discrepancies also exist among the pool of studies that have used tower tasks. This is troubling and invites careful consideration of their methodology. These inconsistencies may partly be understood in terms of sample characteristics: the great variety in ages amongst participants and therefore the role of developmental effects are likely to be important contributory factors. Small sample sizes may also play a role by increasing the likelihood of Type II errors. For example, Corbett et al., (2009) found no evidence for planning impairments in their ASD group but their sample size was small ($n = 18$): it is possible that there was insufficient power to detect a group effect. The ways in which the tower tasks have been administered may also be important to our understanding of the disparities in the literature, as researchers have used a mixture of both computerised and researcher-administered tower tasks. Ozonoff (1995) found that individuals with ASD were less impaired in

cognitive flexibility on the Wisconsin Card Sorting Test (WCST) when it was administered using a computer rather than a researcher in person, highlighting the impact of social cognition on task performance. A similar pattern appears to be emerging among studies investigating planning, with consistent evidence of ASD-related impairments on researcher-administered tower tasks, but generally uncompromised performance on computerised versions being reported (Kenworthy, Yerys, Anthony, & Wallace, 2008). There are therefore a number of factors that may contribute to the discrepancies in the literature; some of which may well play a key role in planning performance in ASD. Carrying out a meta-analysis of the literature will provide an opportunity to address the issues outlined above and therefore advance our understanding of planning in ASD.

1.2.6 Research aims

In light of the mixed evidence on planning impairments in ASD, the present meta-analysis aims to assess whether children with ASD are in fact impaired in this domain. This meta-analysis will be limited to studies that have used tower tasks to assess planning in ASD. The various tower tasks can be thought of as analogous as they assess the same broad cognitive processes and their underlying principles are very similar. It is therefore reasonable to synthesise data from these tasks as a way of reducing the heterogeneity of the data, and therefore the threat to the validity of the meta-analysis (Sharpe, 1997).

This paper will also explore factors that we postulate may account for the disparities seen in the literature. These factors are listed below:

- **Small sample size:** By definition, this meta-analysis will address the issue of small sample size by pooling data from a number of sources.

- Age: This review will explore the effects of participant age and will be limited to studies involving children so as to reduce the heterogeneity between study participants and therefore improve the validity of our findings (Sharpe, 1997).
- Mode of administration: The effects of computerised vs. researcher administration of the tasks will be investigated.
- Study quality: The quality of the methodology for each study will be assessed and examined in terms of its impact on planning performance.

Thus the present meta-analysis aims to contribute to our understanding of planning in ASD and the factors associated with it. Contributing to our understanding of the executive profile in ASD in this way is clinically valuable as it not only provides clues as to the neurocognitive mechanisms implicated in ASD, but also highlights areas of difficulty in ASD which can then be targeted with suitable interventions (Hill, 2004b).

1.3 Method

1.3.1 Data sources and study inclusion

A systematic literature search was carried out using three electronic databases (PSYCinfo, EMBASE and MEDLINE). Search terms related to autism were combined with terms associated with executive functioning, planning, and children (see Figure 1.1 for details). Once duplicate citations were removed, 149 citations remained. The abstracts of these articles were then assessed with the following criteria in mind:

- 1) Planning was being investigated;
- 2) Participants were children with ASD who met diagnostic criteria according to

either the Diagnostic and Statistical Manual of Mental Disorders (3rd edition revised) (DSM-III-R; American Psychiatric Association, 1987), DSM IV (American Psychiatric Association, 1994), DSM 5 (American Psychiatric Association, 2013a), or ICD-10 (International Statistical Classification of Diseases and Related Health Problems – 10th revision; World Health Organization, 1992);

- 3) A typically developing comparison group was included;
- 4) A tower task variant was used (details of these tasks are summarised in Table 1.1);
- 5) Studies were published in a peer-reviewed journal and written in English.

Only studies that appeared to meet these criteria were retrieved for more detailed evaluation. Thirty articles appeared to be appropriate but three of these could not be retrieved as the institution within which this research was conducted did not have access to the relevant journals. The remaining articles were then assessed and only studies adhering to the above eligibility criteria were included. One paper was excluded on the basis that it did not contain information on group means and sample size so effect sizes could not be calculated. This search is summarised in Figure 1.1.

1.3.2 Study quality

Study quality was assessed using the Critical Appraisal Skills Programme (CASP, 2014) checklist for case control studies (see Appendix I for details). This is an 11-item measure that assesses the quality of studies' methodology and findings. It includes two qualitative items and nine quantitative items, therefore yielding a maximum possible score of nine. Where it was unclear from the paper whether the study met certain criteria attempts to contact the author were made. Eleven out of 17 authors were contacted, and six authors failed to respond. In the cases where

evidence could not be obtained, a conservative approach was taken and it was assumed that the studies in question did not meet the particular criteria i.e. a score of zero was awarded.

1.3.3 Planning measures

Although all the studies included in the meta-analysis used tower tasks to assess planning, the dependent variables reported varied. Some authors presented the total number of problems solved in the minimum number of moves, while others reported a score derived from the number of trials required to solve a problem. Other studies reported the number of extra moves required to reach the goal configuration, or the total number of moves needed to complete the task. In addition, some authors included data on the initial thinking time (time between being presented with the target goal state and starting to move the rings or balls) and/or the subsequent thinking time (time taken to complete reach the goal configuration once they have started). In some of the papers that were included, authors had devised their own idiosyncratic dependent score, and others reported a combination of some or all of the above.

Due to the variety and quantity of dependent measures used in these studies, we selected the variable judged best to represent participants' overall planning efficiency and performance: the number of problems solved in the minimum number of moves. This was the most commonly reported measure and was included in 6 of the 17 studies. When this was not reported, the dependent measure that was most related to overall performance and efficiency was used in the meta-analysis. Dependent variables that involved a calculation of planning success in terms of the number of moves required to solve problems were favoured over other measures, such as initial or subsequent thinking time, as these measures were judged to be more

similar to the most frequently reported outcome measure i.e. the number of problems solved in the minimum number of moves ($n = 8$). One of these papers reported performance scores for both the computerised and standard versions of the Tower of London task, as well as a composite of the two. In this case the combined score was used for effect size calculations. The remaining three studies reported total scores without explaining how these were derived. These scores were included in the analysis as they were the only scores reported in these papers. Similar methods have been used in a number of other meta-analyses (e.g. Wykes, Huddy, Cellard, McGurk, & Czobor, 2011). The dependent measures that would be included in the analysis were finalised before effect sizes were calculated so as to reduce bias (Geurts, van den Bergh, & Ruzzano, 2014). The dependent measures that were used for each study in this review are reported in Table 1.2.

1.3.4 Effect size calculation

Standardised effect sizes (Cohen's d ; Cohen, 1977) of planning ability measured by tower tasks were calculated using the difference between the means for the control group and ASD group divided by the pooled standard deviation for the two groups. This is summarized in the following equation (taken from (Wykes et al., 2011):

$$\text{Effect size } (d) = (M_{\text{ASD}} - M_{\text{C}}) / SD_{\text{pooled}}$$

M_{C} indicates the mean for the control group, and SD_{pooled} indicates the pooled standard deviation for the control and ASD groups. Effect sizes were interpreted in line with the guidance by Cohen (1977): small = 0.2; medium = 0.5; large = 0.8. A negative effect size indicates poorer task performance in the ASD group, and a positive effect size indicates that ASD participants performed better than controls. Effect sizes were reversed for those studies where lower scores indicated better

performance and greater planning efficiency. This was the case for three of the studies included in the analysis. The estimated standard error of Cohen's d was calculated using the following equation (taken from Wykes et al., 2011):

$$SE(d) = \sqrt{\left(\frac{n_{ASD} + n_c}{n_{ASD} \times n_c}\right) + \left(\frac{d^2}{2 \times (n_{ASD} + n_c - 2)}\right)}$$

The n_{ASD} and n_c denote the sample size for both the ASD and typically developing control groups (Cooper & Hedges, 1994).

1.3.5 Statistical procedures

It was anticipated that there would be variability at both the study level and the subject level (Lipsey & Wilson, 2001). Therefore a random effects model was used and conducted using STATA software (StataCorp, 2011). The chi-square value of the homogeneity of effects was determined with the Q statistic (Hedges & Olkin, 1985).

Table 1.1**Summary of the tower task variants**

Tower Task Variant	Description
Stockings of Cambridge	Three differently coloured balls are arranged in a specific configuration at the top of a computer screen. Three identical balls are presented at the bottom of the computer screen in a different configuration. The participant is required to match these with the goal set. They are told the minimum number of moves necessary to match the goal configuration and are instructed to use as few moves as possible (Ozonoff et al., 2004).
Tower of Hanoi*	Doughnut-like discs graded in size to form a pyramid-like structure must be moved from one of three identical pegs to another. The participant must follow these rules: only one disc can be moved at a time; if there is more than one disc on a peg, only the top disc can be moved; and discs can only be placed above a larger disc (Ozonoff et al., 2004).
Tower of London*	Three differently coloured balls are arranged on differently sized pegs. The pegs can hold three balls, two balls, or just one ball. A graded set of problems requiring a minimum of 2, 3, 4, or 5 moves is presented (Ozonoff et al., 2004).
NEPSY Tower	Three differently coloured balls are arranged in a specific configuration and displayed on a picture board. The participant is required to rearrange three identical coloured balls arranged on three pegs to match the goal set. Participants must reach this configuration in a certain number of moves and must follow these rules: only one ball can be moved at a time; balls must be moved directly from one peg to another (Korkman et al., 1998).
D-KEFS Tower	Between two and six differently sized disks are arranged on three pegs. The participant is required to rearrange the same number of disks to match the goal configuration. Participants are required to do this in the minimum number of moves and must follow these rules: bigger disks cannot be placed on smaller disks; only one disk can be moved at a time; all disks must be placed on a peg (Delis et al., 2001).

**The exact variant used varies slightly between studies in terms of the number of puzzles presented, whether they were graded in difficulty, and the exact method of administration. However, the underlying principles behind these tasks are very similar.*

Figure 1.1

Literature search strategy

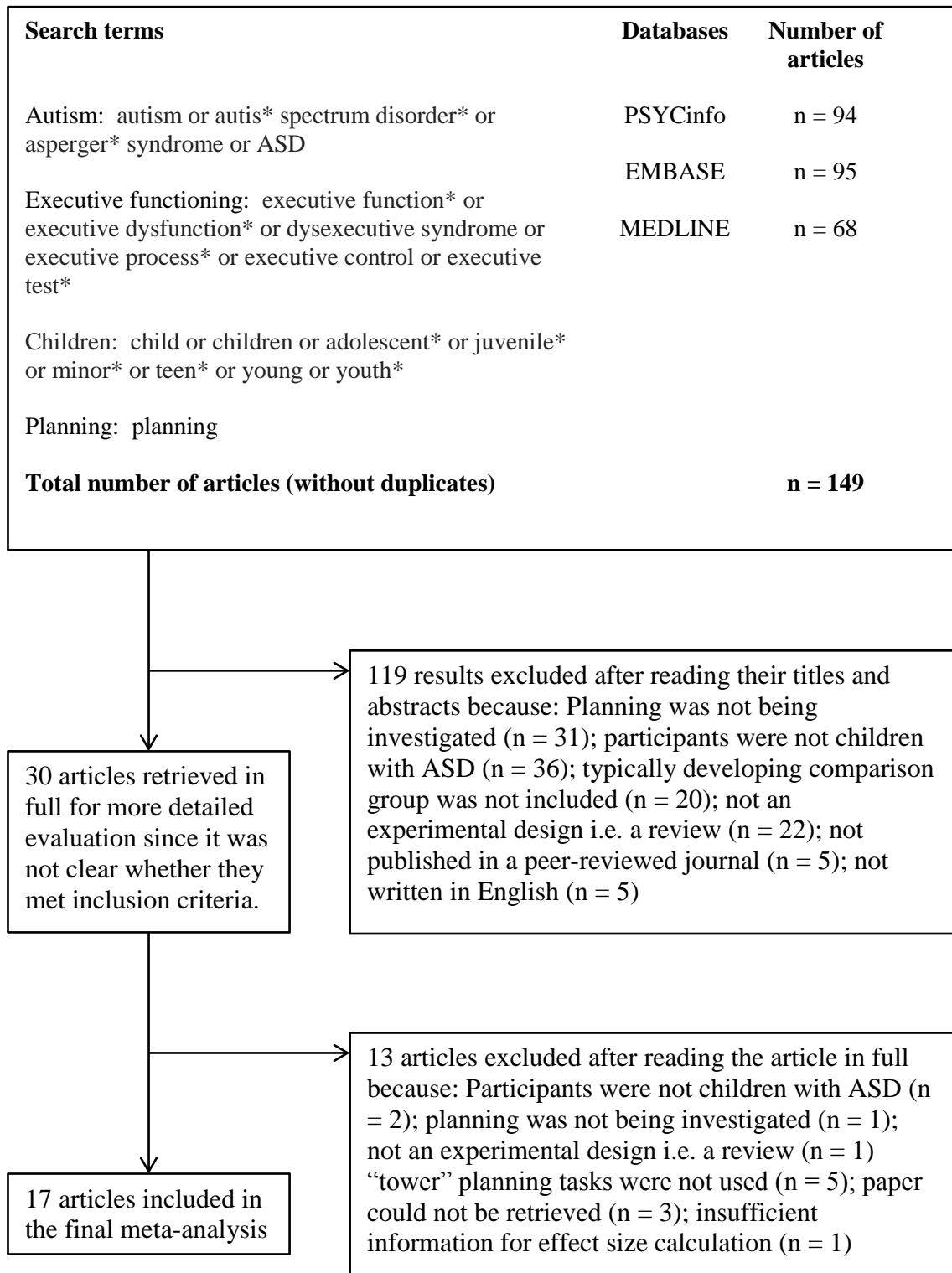


Table 1.2**Summary of the dependent measures used for the purposes of the meta-analysis**

Study	Dependent measure used for meta-analysis
Corbett et al. (2008)	Total number of problems solved in the minimum number of moves
Geurts et al. (2004)	Total score calculated by assigning points based on the number of trials required to solve a problem
Goldberg et al. (2005)	Total number of problems solved in the minimum number of moves
Happe et al. (2006)	Number of extra moves required to complete problems requiring 4 and 5 move solutions (NB. Regression co-efficients reported)
Joseph et al. (2005)	Total number of problems solved in the minimum number of moves
Kimhi et al. (2014)	Total score calculated by assigning points based on the number of trials required to solve a problem
Landa & Goldberg (2005)	Total score calculated by assigning points based on the number of trials required to solve a problem
Ozonoff & Jensen (1999)	A total score is reported but it is not explained how this was calculated
Panerai et al. (2014)	A total score is reported but it is not explained how this was calculated
Pellicano (2007)	Total number of problems solved in the minimum number of moves
Pellicano (2010)	Total number of problems solved in the minimum number of moves
Robinson et al. (2009)	Number of moves used to solve problems
Semrud-Klikeman et al. (2010)	A total achievement score is reported but it is not explained how this was calculated
Sinzig et al. (2008)	Total number of problems solved in the minimum number of moves
Van Eylen et al. (2015)	Move accuracy ratio (actual number of moves divided by the number of minimally required moves)
Verte et al. (2005)	Total score calculated by assigning points based on the number of trials required to solve a problem
Williams & Jarrod (2013)	Total number of moves used to solve problems

1.4 Results**1.4.1 Corpus of studies**

The literature search yielded 17 studies that fulfilled all the inclusion criteria set out above. The studies that were included in the final meta-analysis are summarised in Table 1.3.

1.4.2 Study quality

Table 1.4 summarises the study quality scores for each study. The mean score was 6.88 (SD=0.93) out of a maximum of nine. Problems with study quality included small sample size, possible selection biases, difficulties ruling out possibly confounding factors, and doubts about the generalisability of the findings.

1.4.3 Meta-analysis results

Figure 1.2 illustrates a forest plot for performance in tower planning tasks, showing an overall significant negative effect of ASD of -0.65 (95% confidence interval [CI]=(-0.84)-(-0.46)). This is a medium to large effect (Cohen, 1977). Only two studies found that participants with ASD performed better on tower tasks than typically developing children.

Heterogeneity between studies was significant ($Q = 36.57$, $df = 16$, $p = 0.002$) meaning that the variability across the effect sizes is greater than would be expected from sampling error alone (Lipsey & Wilson, 2001).

1.4.4 Publication bias

Publication bias was assessed visually using a funnel plot (see Figure 1.3). This did not appear to show a publication bias but Kendall's method (Begg & Mazumdar, 1994) was used to confirm this. The results did not provide evidence for a publication bias ($z = -1.48$, $p = 0.14$).

Table 1.3

Summary of studies included in the meta-analysis: pale grey shading indicates that task was administered in a computerised format; dark grey shading indicates that both computersied and non-computerised versions were used.

Study	Tower task	N (ASD, TD)	Age (mean years; range)	Male:female	Diagnosis	IQ matching
Corbett et al. (2008)	SoC ¹	36 (18,18)	ASD = 9.4 TD = 9.6 Range = 7-12	ASD = 17:1 TD = 12:6	High functioning ASD	Not matched for but used IQ as a covariate
Geurts et al. (2004)	ToL ²	82 (41, 41)	ASD = 9.4 TD = 9.1 Range = 6-12	All male	High functioning autism	Not matched for full scale IQ but used IQ as a covariate
Goldberg et al. (2005)	SoC	49 (17, 32)	ASD = 10.3 TD = 10.4 Range = 8-12	ASD = 13:4 TD = 21:11	High functioning autism	Not matched for full scale IQ but used IQ as a covariate
Happé et al. (2006)	SoC	60 (30, 30)	ASD = 11.1 TD = 11.3 Range = 8-16	All male	High functioning autistic disorder (n = 6) or Asperger disorder (n = 26)	Full scale IQ matched
Joseph et al. (2005)	NT ³	68 (37, 31)	ASD = 7.9 TD = 8.3 Range = 5 - 11	All male	Autism	Yes, for verbal and non-verbal IQ*
Kimhi et al. (2014)	ToL	59 (29,30)	ASD = 5.0 TD = 4.6 Range = 3-6	ASD = 25:4 TD = 26:4	Intellectually able with ASD	IQ matched
Landa & Goldberg (2005)	SoC	38 (19, 19)	ASD = 11.0 TD = 11.0 Range = 7-17	Not known but report that groups are matched	High functioning autism	Full scale IQ matched
Ozonoff & Jensen (1999)	ToH ⁴	69 (40, 29)	ASD = 12.6 TD = 12.1	Not reported	Autistic disorder (FSIQ 70 and	Not matched for full scale IQ but used as a

				Range = 6-18		above)	covariate
Panerai et al. (2014)	ToL	61 (27, 34)	ASD = 9.8 TD = 11.3 Range = unknown	ASD = 22:5 TD = 25:9	High-functioning ASD (n = 11), ASD with borderline intellectual functioning (n = 8), ASD with mild intellectual disability (n = 8)		IQ matched
Pellicano (2007)	ToL	70 (30, 40)	ASD = 5.6 TD = 5.5 Range = 4-7	ASD = 25:5 TD = 31:9	ASD (FSIQ 80 and above)		Matched for verbal and non-verbal IQ
Pellicano (2010)	ToL	72 (37, 31)	ASD = 5.7 TD = 5.4 Range = 4-7	ASD = 33:4 TD = 25:6	ASD (FSIQ 80 and above)		Matched for verbal and non-verbal IQ
Robinson et al. (2009)	ToL	108 (54, 54)	ASD = 12.5 TD = 12.1 Range = 8-17	ASD = 42:12 TD = 42:12	High functioning autism or Asperger Syndrome		Full scale IQ matched
Semrud-Klikeman et al. (2010)	D-KT ⁵	47 (15, 32)	ASD = 10.6 TD = 9.8 Range = 9-16	ASD = 8:7 TD 23:9	Asperger disorder		Full scale IQ matched
Sinzig et al. (2008)	SoC	40 (20, 20)	ASD = 14.3 TD = 13.1 Range = 6-18	ASD = 16:4 TD = 14:6	High functioning autism (n=5) or Asperger Syndrome (15)		IQ Matched
Van Eylen et al. (2015)	D-KT	100 (50, 50)	ASD = 12.2 TD = 12.5 Range = 8-18	ASD = 30:20 TD = 30:20	ASD (FSIQ 70 and above)		Full scale IQ matched
Verte et al. (2005)	ToL	108 (61, 47)	ASD = 9.1 TD = 9.4 Range = 6-13	ASD = 57:4 TD = 40:7	High functioning autism		Not matched for full scale IQ but used IQ as a covariate
Williams & Jarrold (2013)	ToL	43 (21, 22)	ASD = 10.45 TD = 10.61 Range = unknown	Not reported	ASD		Verbal IQ and performance IQ matched

Note. Diagnoses have been described exactly as presented in the text. ASD = Autism Spectrum Disorder, TD = Typically developing controls.

¹SoC = *Stockings of Cambridge*

²ToL = *Tower of London*

³NT = *NEPSY Tower*

⁴ToH = *Tower of Hanoi*

⁵D-KT = *D-KEFS Tower*

*Participant IQs were not all within the normal range i.e. some below 70

Table 1.4

Critical Appraisal Skills Programme (CASP, 2014) checklist scores

Study	CASP score (out of a maximum of nine)
Corbett et al. (2008)	6
Geurts et al. (2004)	6
Goldberg et al. (2005)	6
Happe et al. (2006)	6
Joseph et al. (2005)	8
Kimhi et al. (2014)	8
Landa & Goldberg (2005)	6
Ozonoff & Jensen (1999)	6
Panerai et al. (2014)	8
Pellicano (2007)	8
Pellicano (2010)	8
Robinson et al. (2009)	8
Semrud-Klikeman et al. (2010)	6
Sinzig et al. (2008)	6
Van Eylen et al. (2015)	7
Verte et al. (2005)	7
Williams & Jarrold (2013)	7

Figure 1.2

A forest plot of planning performance effect sizes and confidence intervals for each study included in the meta-analysis

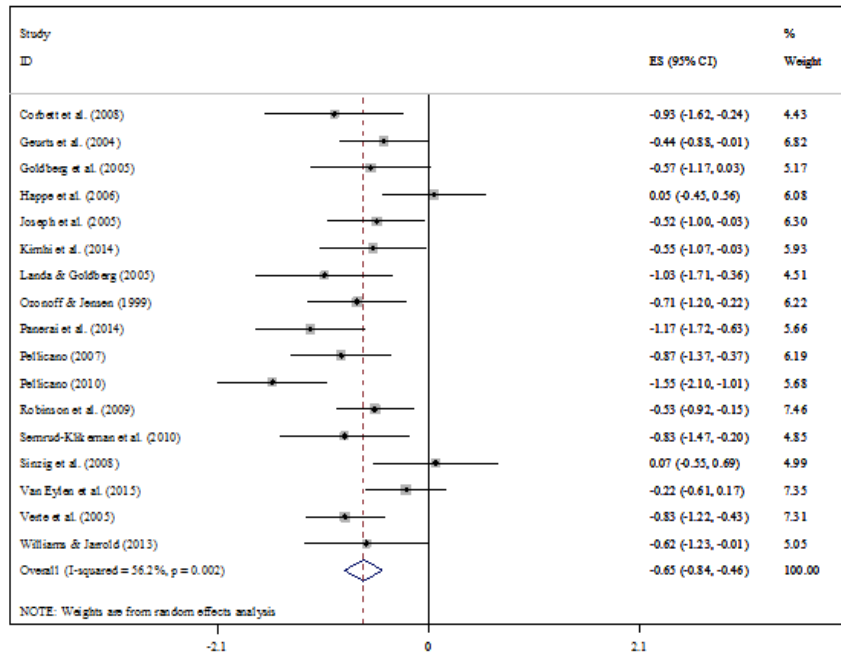
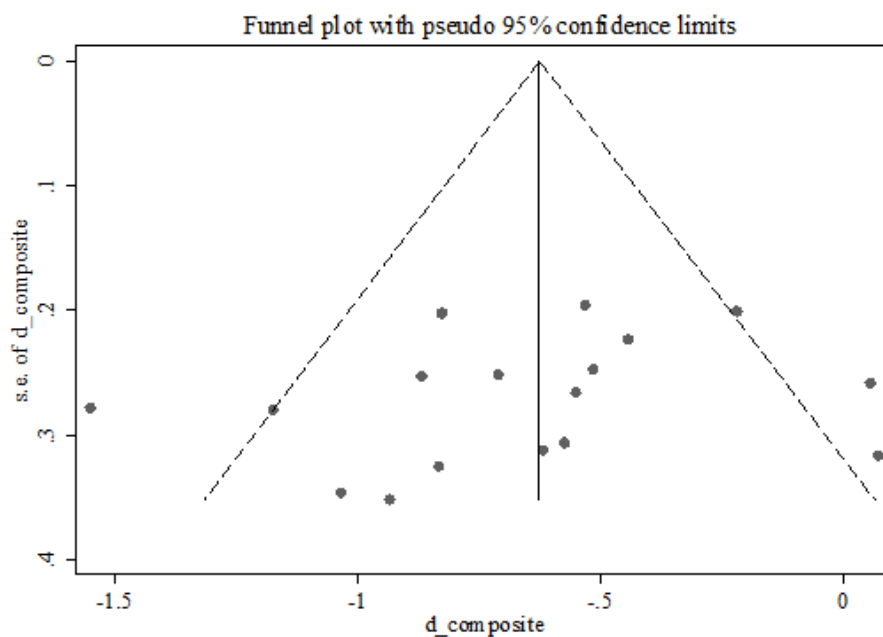


Figure 1.3

A funnel plot to explore for publication bias



1.4.5 Factors affecting planning ability

Heterogeneity among the studies was explored using categorical analysis. This approach has been taken by other authors where fewer than 20 effect sizes are available and there is therefore insufficient statistical power to carry out a meaningful meta-regression (e.g. Wykes et al., 2011).

1.4.5.1 Mode of administration. Previous findings for both planning and other domains of executive functioning, such as flexibility, (Ozonoff, 1995), indicate that ASD participants perform better when tasks are administered using a computer as opposed to a researcher. It may be that socially-administered tasks present a particular challenge for ASD participants who have both social cognition and motivational deficits (Kenworthy et al., 2008). The studies included in the meta-analysis were therefore grouped according to whether the tower tasks were administered manually by a researcher or by computer. For studies where computerised versions of the tasks were used there was a trend effect ($n = 6$): effect size = -0.455, 95% CI = -0.916 to 0.006. For studies where the tasks were administered by a researcher, there was a significant effect ($n = 12$): effect size = -0.712, 95% CI = -0.909 to -0.515.

1.4.5.2 Participant age. It was hypothesised that age-related improvements in planning ability might be observed, as suggested by Happé et al. (2006) and Ozonoff & Jenson (1999). In order to explore the effect of age, studies were divided into two groups according to the mean age of their participants. The first group comprised studies where the mean age was less than 11 ($n = 10$). The second group included children aged 11 and over ($n = 7$). There were significant effects for both groups (mean age under 11 [$n = 10$]: effect size = -0.793, 95% CI = -1.016 to -0.570; mean age 11 and over [$n = 7$]: effect size = -0.457, 95% CI = -0.746 to -0.168).

1.4.5.3 Study quality. In order to explore the effect of study quality, the studies were split into two groups: those that scored seven and above using the CASP scale and those that scored below seven. There were significant effects for both groups (CASP score seven and above [n = 9]: effect size = -0.741, 95% CI = -0.995 to -0.487; CASP score below seven [n = 8]: effect size = -0.523, 95% CI = -0.805 to -0.241).

1.5 Discussion

The central aim of this meta-analysis was to determine whether planning deficits are present in children with ASD. A variety of tasks have been used to measure planning in ASD in the literature, but this review focused on studies that used tower tasks in order to reduce heterogeneity between studies and therefore the threat to the validity of the meta-analysis (Sharpe, 1997). Tower tasks have the additional benefit of having been extensively used to measure planning in ASD, as well as other in populations (e.g. Owen, Downes, Sahakian, Polkey, & Robbins, 1990), meaning that there was a substantial body of data to include in the review and that it would produce results that could be readily understood by other researchers. The findings reported here demonstrate a medium to large effect for the difference in tower task performance between typically developing children and children with ASD. The direction of the effect suggests that children with ASD are typically outperformed by healthy controls on these tasks. This confirms the findings from previous non-systematic, narrative reviews (Hill, 2004a, 2004b). In addition there is no evidence of a publication bias which would present a further threat to the validity of the findings (Sharpe, 1997).

This review additionally investigated factors that we proposed might be associated with planning performance in ASD. Due to the small number of available effect sizes, the effect of these factors on task performance was explored using categorical analyses, whereby studies were grouped according to variables of interest (mode of administration, age, study quality,) and pooled effect sizes were calculated separately for each group. We did not formally investigate the relationships between variables of interest and effect size, as we had insufficient power for such analyses (Wykes et al., 2011).

When studies were grouped according to whether tasks were administered by a researcher or by computer, we found that there was a significant group effect for the researcher administered group, but not for the computer administered group. This provides tentative support for the idea that planning performance is moderated by the social demands of the tasks used to assess it, though it is worth noting that there may have been insufficient power to detect an effect for planning difficulties when the task was administered electronically due to the relatively small subset of studies that took this approach. It may be that in a task involving a seemingly arbitrary set of rules, individuals with ASD are not conscious of the social expectation to obey these rules and therefore follow their own preferences instead. White et al. (2009) hypothesise that Theory of Mind may therefore moderate executive performance in ASD (Joseph & Tager-Flusberg, 2004; Ozonoff et al., 1991; Pellicano, 2007). Further research into both planning and other executive domains could provide valuable clarification in terms of the impact of social demands on executive performance, as well as the mechanisms involved. Whether or not mentalising moderates executive dysfunction, if socially demanding tasks prove to be more challenging for individuals with ASD, there are important clinical

implications. For example, it may be advisable for interventions targeting difficulties to focus on supporting ASD individuals in contexts where social demands are high, perhaps by reducing these demands as much as possible through the use of computerised aids. If social demands act as a moderator, there are also implications for the way future research might be approached. For example, researchers may opt to limit the social or mentalising aspects of the measures they employ in order to assess executive functioning in its “purest” form. Alternatively, they might choose to increase the ecological validity of their research by deliberately assessing executive functioning in the context of social demands.

The effect of age on planning performance was also investigated using categorical analyses. Other researchers have found age-related changes in planning ability amongst their participants (e.g. Happé et al., 2006), as well as in other domains of executive functioning. For example, one study found improvements in inhibition over time during childhood (Huizenga, Ingmar, & Conor, 2011). Similarly, others have reported improvements in cognitive flexibility and ToL performance in typically developing children up to early adolescence (Best, Miller, & Jones, 2009). Interestingly, we did not find evidence to support such a pattern, since a significant planning deficit in ASD was found for both our older and younger groups. Nevertheless, as above, these can only be thought of as tentative findings, which will need to be further investigated in well-powered, longitudinal studies. Our lack of evidence for an age effect on planning problems in ASD may seem surprising in the context of the above research findings but other studies have reported similar null findings in other areas of executive functioning, such as inhibition (Geurts et al., 2014).

Meta-analysis findings are only as good as the studies that are included in the meta-analysis (Lipsey & Wilson, 2001). The methodological quality of the studies included in this meta-analysis was therefore carefully assessed using the CASP scale and the influence of study quality on the findings was subsequently assessed using categorical analysis. Study quality was high over all and all the studies included in the final analysis scored six and above out of nine total points. The analysis indicated that study quality had no effect on planning performance, though as all of the studies included in this review scored relatively highly this is perhaps unsurprising. The same precautions with regard to interpreting these findings apply due to the limitations of the method used.

Only studies using tower planning tasks were included in this meta-analysis with the aim of reducing heterogeneity between the study methodologies and therefore increasing the validity of the findings (Lipsey & Wilson, 2001). However, some wider issues associated with the use of tower tasks merit consideration. Firstly, although the tower tasks have good face validity, concerns have been raised about their construct validity and psychometric properties (e.g. Kafer & Hunter, 1997). These tasks arguably tap a number of cognitive functions including planning, inhibition, procedural learning, explicit reasoning, working memory, visuospatial skills, and rule-following. It is therefore difficult to draw definitive conclusions from tower task data about planning alone (Ozonoff et al., 2004) and further investigation of the specific constructs involved is required. Perhaps even more important than their construct validity is the tasks' ecological validity. Tower tasks are of an abstract and highly structured nature and therefore lack ecological validity in terms of their ability to emulate real-world planning (de Jager et al., 2014). The ecological validity of neuropsychological measures is increasingly prioritised (e.g. Burgess et

al., 2006) as it not only has a direct impact on their clinical usefulness, but is also thought to increase measure sensitivity (Hill & Bird, 2006). This lack of ecological validity is a significant limitation and unfortunately seems to extend to the vast majority of the measures of planning that appear in the ASD literature (Kenworthy et al., 2008). For example, the mazes tasks that have been used to assess planning bear little resemblance to real-world planning scenarios, requiring participants to plan, to time, their route through a diagram of a maze. Prioritising ecological validity in both task and research design would be invaluable in advancing our understanding of planning abilities, and therefore the executive profile, in ASD.

The high incidence of comorbidity in ASD, particularly with regard to ADHD and obsessive compulsive disorder (Leyfer et al., 2006), is an important issue to consider in relation to these findings. These disorders are also associated with impaired executive functioning and failure to tease these disorders apart in empirical research confuses our understanding of the executive profile in ASD. For example, some researchers have found an additive effect for comorbidity, where children with ASD and ADHD diagnoses had greater inhibition problems than those with an ASD diagnosis alone (Sinzig et al., 2008). With respect to this meta-analysis, efforts were made to gain clarity on the specific clinical characteristics of the ASD groups for each study. However, it was not always possible to get a definitive answer and an estimated seven studies included in this review are likely to have included ASD children with comorbid diagnoses. This is slightly problematic in terms of moving towards defining the executive profile for ASD as we cannot rule out the confounding contribution of other disorders for these few studies. However, as comorbidity is so common amongst ASD individuals the clinical relevance of this research is arguably uncompromised in the sense that the findings reported here are

probably generalisable to the broader ASD population.

As many as 25-40% of ASD cases are affected by comorbid intellectual disability (Chakrabarti & Fombonne, 2001). As the majority of the studies included in this review focused on individuals with an IQ above 70 ($n = 14$), the generalisability of our findings across the entirety of the autism spectrum may have been jeopardised. Limiting our sample to children may have posed a similar threat as obviously people of all ages are affected by ASD as it is a lifelong developmental condition. However, it was deemed that protecting the validity of the meta-analysis should be prioritised.

In conclusion, this review found evidence for impaired performance in tower planning tasks amongst children with ASD. We found no effect for age or study quality on task performance. However, our findings offer tentative support for the idea that the social demands of laboratory tasks may affect executive performance in ASD and therefore influence how accurately executive dysfunction is assessed. This is interesting in the context of White et al.'s (2009) hypothesised moderation of executive functioning by Theory of Mind in ASD. The relatively small numbers of intellectually impaired individuals present in the sample may limit the generalisability of the findings, as might the limited age range of the participants. In terms of other sample characteristics, it is likely that a number of the participants had comorbid disorders. This may confound our results but equally makes them more clinically relevant as comorbidity within ASD is so common.

Overall, this meta-analysis contributes to our understanding of the executive profile in ASD as it provides evidence for impaired performance on tower tasks. However, it is not possible to draw definitive conclusions about planning per se in the context of the task limitations discussed above. Perhaps most concerning is the

abstract and structured nature of the tower tasks themselves, as well as that of their counterparts, and therefore their poor semblance to real-world scenarios. This undermines the tasks' clinical relevance, thus highlighting the value in prioritising ecological validity in both task and research design moving forward. This shift will allow us to move towards a far more meaningful characterisation of the executive functioning profile in ASD.

1.6 References

- American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders third edition revised. *Washington, DC: The Association.*
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders fourth edition. *Washington DC: American Psychiatric Association, 69–81.*
- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders fifth edition. *Arlington: American Psychiatric Publishing, 280.*
- Baird, G., Charman, T., Baron-Cohen, S., Cox, a, Swettenham, J., Wheelwright, S., & Drew, a. (2000). A screening instrument for autism at 18 months of age: a 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 39(6), 694–702.*
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368(9531), 210–215.*
- Barkley, R. A. (1998). Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment (2nd ed.). *New York, NY, US: Guilford Press, 628.*
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50(4), 1088–1101.*
- Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review.*
- 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, 194–209.
- CASP. (2014). CASP Checklists.
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *JAMA*, 285(24), 3093–9.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc, 474.
- Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis. The handbook of research synthesis*.
- Corbett, B., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research*, 166(2–3), 210–222.
- de Jager, C. a., Dye, L., de Bruin, E. a., Butler, L., Fletcher, J., Lamport, D. J., ... Wesnes, K. (2014). Criteria for validation and selection of cognitive tests for investigating the effects of foods and nutrients. *Nutrition Reviews*, 72(3), 162–179.
- Delis, D., Kaplan, E., & Kramer, J. (2001). Delis-Kaplan executive function system (D-KEFS). *Canadian Journal of School Psychology*, 20(1–2), 117–128.
- Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3(3), 271–290.
- Geurts, H. M., van den Bergh, S. F. W. M., & Ruzzano, L. (2014). Prepotent response inhibition and interference control in autism spectrum disorders: Two meta-analyses. *Autism Research*, 7(4), 407–420.
- Goldberg, M. C., Mostofsky, S. H., Cutting, L. E., Mahone, E. M., Astor, B. C.,

- Denckla, M. B., & Landa, R. J. (2005). Subtle executive impairment in children with autism and children with ADHD. *Journal of Autism and Developmental Disorders, 35*(3), 279–293.
- Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61*(1), 25–39.
- Hedges, L. V, & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, Academic Press.
- Hill, E. L. (2004a). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24*(2), 189–233.
- Hill, E. L. (2004b). Executive Dysfunction In Autism. *Trends in Cognitive Sciences, 8*(1), 26–32.
- Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia, 44*(14), 2822–2835.
- Hughes, C. (1996). Brief report: Planning problems in autism at the level of motor control. *Journal of Autism and Developmental Disorders, 26*(1), 99–107.
- Huizenga, H. M., Ingmar, V., & Connor, D. (2011). Testing overall and moderator effects in random effects meta-regression. *British Journal of Mathematical & Statistical Psychology, 64*(1), 1–19.
- Joseph, R. M., McGrath, L. M., & Tager-Flusberg, H. (2005). Executive Dysfunction and Its Relation to Language Ability in Verbal School-Age Children With Autism. *Developmental Neuropsychology, 27*(3), 425–458.
- Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and

- executive functions to symptom type and severity in children with autism. *Dev Psychopathol*, 16(1), 137–155.
- Kafer, K. L., & Hunter, M. (1997). On testing the face validity of planning/problem-solving tasks in a normal population. *Journal of the International Neuropsychological Society*, 3(2), 108–119.
- Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding Executive Control in Autism Spectrum Disorders in the Lab and in the Real World. *Neuropsychology Review*, 18(4), 320–338.
- Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. Psychological Corporation.
- Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., ... Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, 36(7), 849–861.
- Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. *Applied Social Research Methods Series*, 49, 264.
- Milner, B. (1965). Visually-guided maze learning in man: Effects of bilateral hippocampal, bilateral frontal, and unilateral cerebral lesions. *Neuropsychologia*, 3(4), 317–338.
- Nydén, A., Hagberg, B., Goussé, V., & Rastam, M. (2011). A cognitive endophenotype of autism in families with multiple incidence. *Research in Autism Spectrum Disorders*, 5(1), 191–200.
- Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, 28(10), 1021–1034.

- Ozonoff, S. (1995). Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism. *Neuropsychology*, 9(4), 491–500.
- Ozonoff, S., Cook, I., Coon, H., Dawson, G., Joseph, R. M., Klin, A., ... Wrathall, D. (2004). Performance on Cambridge Neuropsychological Test Automated Battery Subtests Sensitive to Frontal Lobe Function in People with Autistic Disorder: Evidence from the Collaborative Programs of Excellence in Autism Network. *Journal of Autism and Developmental Disorders*, 34(2), 139–150.
- Ozonoff, S., & Jenson, J. (1999). Brief Report: Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders*, 29(2), 171–177.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 32(7), 1081–1105.
- Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: clues to developmental primacy. *Developmental Psychology*, 43(4), 974–990.
- Pennington, B. F., & Ozonoff, S. (1996). Executive Functions and Developmental Psychopathology. *Journal of Child Psychology and Psychiatry*, 37(1), 51–87.
- Prior, M., & Hoffmann, W. (1990). Brief report: Neuropsychological testing of autistic children through an exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders*, 20(4), 581–590.
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, 27(2), 224–260.
- Semrud-Clikeman, M., Walkowiak, J., Wilkinson, A., & Butcher, B. (2010). Executive functioning in children with asperger syndrome, ADHD-combined

- type, ADHD-predominately inattentive type, and controls. *Journal of Autism and Developmental Disorders*, 40(8), 1017–1027.
- Sergeant, J. a., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? *Behavioural Brain Research*, 130(1–2), 3–28.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, 17(8), 881–901.
- Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child and Adolescent Psychiatry and Mental Health*, 2(1), 4.
- StataCorp. (2011). Stata Statistical Software: Release 12.
- Taylor, A. E., Saint-Cyr, J. A., & Lang, A. E. (1986). Frontal lobe dysfunction in Parkinson's disease. *Brain*, 109, 845–883.
- Van Eylen, L., Boets, B., Steyaert, J., Wagemans, J., & Noens, I. (2015). Executive functioning in autism spectrum disorders: influence of task and sample characteristics and relation to symptom severity. *European Child & Adolescent Psychiatry*.
- Weinberger, D. R., & Gallhofer, B. (1997). Cognitive function in schizophrenia. *International Clinical Psychopharmacology*, 12 Suppl 4(1), S29–S36.
- White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on “open-ended” executive function tests in Autism. *Autism Research*, 2(3), 138–147.
- World Health Organization. (1992). The ICD-10 Classification of Mental and Behavioural Disorders. *International Classification*, 10, 1–267.

Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A Meta-Analysis of Cognitive Remediation for Schizophrenia: Methodology and Effect Sizes, (May), 472–485.

Part 2: Empirical Research Paper

**Developing and Piloting a New Measure of Executive Functioning for Children
with Autism Spectrum Disorder (ASD)**

2.1 Abstract

Aims: To develop and pilot the Ecologically-Valid Test of Executive Dysfunction (Eco-TED) for children with Autism Spectrum Disorder (ASD)

Method: Seven different Eco-TED tasks were developed as part of a joint project. The current study is concerned with three of these: the Consent Form Test and the School Bag and Lego Tasks. Twenty children with ASD and 20 age and IQ-matched controls completed the Eco-TED, as well as two already-developed executive function tasks: the Zoo Map and Six Part Tests from the Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C). Participants were between 8 and 13 years old and had an IQ > 70. A variety of parent-report questionnaires were also administered. Fourteen participants repeated the Eco-TED approximately one month later to assess test-retest reliability. The data were analysed using both raw and empirically derived deficit scores. Individual variability in performance was explored, as well as correlations between outcome measures and retest scores.

Results: No group differences were found when raw scores were analysed. Using deficit scores, group differences were found for the School Bag and Lego Tasks: composite scores indicated that ASD participants were more impaired than controls. Test-retest correlations were low and Eco-TED scores did not generally correlate with BADS-C scores, age or IQ.

Conclusion: Significant differences in impairment were found for the School Bag and Lego Tasks using a deficit scoring method. It seems likely that at least planning, prospective memory, working memory, attention, inhibition and interference control are implicated. The extent to which the Eco-TED correlates with everyday outcome variables, its criterion validity and test-retest reliability all require further investigation.

2.2 Introduction

Autism Spectrum Disorder (ASD) is present in at least 1% of the population and affects three times as many males as females (Baird et al., 2006). It is a lifelong neurodevelopmental condition and is defined by the Diagnostic and Statistical Manual of Mental Disorders fifth edition (DSM-5; American Psychiatric Association, 2013) as a dimensional rather than categorical disorder. Diagnostic criteria comprise social interaction difficulties; specifically, deficits in socio-emotional reciprocity, verbal and non-verbal communication, and difficulties in forming and maintaining relationships. The DSM-5 also describes a second set of symptoms, including repetitive and stereotyped behaviours and speech, excessive adherence to routines and rituals, highly fixated and restricted interests, and sensory abnormalities. Individuals must display at least two characteristics from this second domain to meet the diagnostic threshold. Presentations of diagnosed individuals vary greatly in terms of their behaviour, the severity of their symptoms (Nydén et al., 2011) and their cognitive abilities. In fact, the variability both within and between individuals with ASD in terms of their cognitive profile is considered by some to be a defining feature of this group (Towgood, Meuwese, Gilbert, Turner, & Burgess, 2009).

Executive functioning comprises a variety of skills and functions including planning, flexibility, inhibition, shifting set, impulse control, working memory, generativity, and initiation and monitoring of action. These domains are all thought to be involved in the overarching control of behaviour in order to attain a goal (Duncan, 1986). Executive functioning has been shown by neuroimaging studies to be associated with the frontal lobes (Monchi, Petrides, Strafella, Worsley, & Doyon, 2006; Stuss & Alexander, 2000; Stuss et al., 2009), and patients with acquired frontal

lobe damage are typically impaired in tasks which aim to measure these functions (White et al., 2009). Behavioural similarities observed between frontal lobe patients and individuals with ASD led to the development of a theory of executive dysfunction in ASD (e.g. Ozonoff, Pennington, & Rogers, 1991). This is now one of the prominent cognitive theories of ASD, seeking to explain both the social and non-social features of the condition in terms of executive dysfunction. For example, its proponents argue that cognitive flexibility deficits can account for both perseveration during daily activities and difficulties in adapting perspective during social interactions (Geurts, Corbett, & Solomon, 2009). This theory also has explanatory power for some symptoms not adequately explained by the other influential cognitive accounts of ASD. For example, executive dysfunction offers a plausible explanation for the need for sameness and impeded impulse control in ASD (Rajendran & Mitchell, 2007), both of which cannot be explained satisfactorily by the theory of mind deficit hypothesis (Happé, 1994); weak central coherence accounts (Frith, 1989, 2003); and the extreme male brain theory (Baron-Cohen, 2002).

Executive functioning in ASD has been extensively researched. While some studies have reported impaired executive performance in their ASD groups, findings in this field are mixed. In a review of studies of executive functioning in ASD, Hill (2004a) found evidence for significant planning and cognitive flexibility deficits, supporting earlier findings (B F Pennington & Ozonoff, 1996; Sergeant et al., 2002). The review also provided some evidence for impairments in other executive domains; namely response inhibition and generativity. However the reported effects were generally fairly modest in size, and the review also highlighted numerous disparities between findings across the various age groups and functions

investigated. More recent reviews present a similarly mixed picture. For example, Geurts et al. (2009) argue that ‘bottom-up’ factors, such as stress, are likely to contribute to poor task performance and therefore question the validity of the assumed link between observed behavioural rigidity in research tasks and the cognitive flexibility they aim to measure. They conclude that there is no consistent evidence for a deficit in cognitive flexibility in ASD, while Russo et al. (2007) maintain, like Hill (2004a), that there is clear evidence for such an impairment. Furthermore, many researchers have been unable to detect executive deficits in their ASD participants (Baron-Cohen, Wheelwright, Stone, & Rutherford, 1999; Hill & Russell, 2002; Minshew, Goldstein, Muenz, & Payton, 1992; Russell & Hill, 2001). Such findings contrast with anecdotal evidence from clinicians, teachers and families of individuals with ASD: those close to individuals with ASD typically identify a number of substantial executive deficits that cause daily problems for them (Kenworthy et al., 2008).

The disparities described above have raised questions about the methodological approaches typically employed in this field. In particular, they have invited careful consideration of the suitability of existing measures of executive functioning in assessing this construct in ASD. Many of the tasks so often used in ASD research, such as the Wisconsin Card Sorting Test (WCST; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), were originally created for use in areas far removed from ASD, such as the assessment of acquired brain injury. They are arguably now often used inappropriately, or at least in a way that is not optimal for their purpose (Burgess et al., 2006). Research has highlighted the difficulties this presents. For example, individuals with ASD have been shown to be less impaired in flexibility on the WCST when it is administered using a computer rather than a researcher in

person (S Ozonoff, 1995). Therefore, more traditional, socially administered neuropsychological tasks, not having been designed for individuals with ASD, may fail to account for the impact of social cognition and motivational deficits on task performance (Kenworthy et al., 2008).

A further methodological issue that has been raised is the ‘ecological validity’ of current measures of executive functioning. This is increasingly prioritised and relates to the extent to which a measure can accurately reflect and predict difficulties experienced in real-life situations (Kenworthy et al., 2008). Arguably, the more a task emulates the demands of real-life scenarios, and thus the more ecologically valid it is, the more clinically useful it may be (Burgess et al., 2006). So-called ‘classic’ executive functioning measures, so often used in this field, can be critiqued for lacking ecological validity, such that they may fail to capture executive deficits amongst ASD participants. Hill & Bird (2006) compared a battery of ‘classic’ tests of executive functioning, including the Modified Wisconsin Card Sorting Test (MWCST; Nelson, 1976), with two newer, more ecologically valid tests: the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Burgess, Alderman, Wilson, Evans, & Emslie, 1996) and the Hayling test (Burgess & Shallice, 1997), both of which have evidenced ecological validity (Odhuba, van den Broek, & Johns, 2005). Only the BADS and Hayling test detected any group differences in performance between adults with Asperger syndrome and controls. Executive functioning deficits are therefore “best observed through using more recent, ecologically valid tests” (Hill & Bird, 2006, p. 2822). This view is now quite widely shared (e.g. Kenworthy et al., 2008), but unfortunately many existing ecologically valid tests of executive functioning have their own limitations.

Kenworthy et al. (2008) provide an elaborated framework for appraising ecological validity. They propose that comprehensive assessment of ecological validity involves consideration of verisimilitude and veridicality. Verisimilitude describes the tasks resemblance to real-life demands, while veridicality addresses the extent to which a test correlates with everyday outcome variables. With this in mind, Kenworthy et al. (2008) discuss the BADS, which they describe as “the most commonly-used and comprehensively ecologically-valid” (p. 327) measure of executive functioning. They question the veridicality of this test, pointing out that the BADS does not always correlate with measures of everyday functioning. For example, Wood & Lioffi (2006) found that the Zoo Map and Key Search subtests of the BADS did not correlate with scores from the Dysexecutive Questionnaire (DEX; Burgess et al., 1996); a questionnaire-based measure which forms part of the extended BADS battery and uses informant reports to assess everyday functioning. The child version of the BADS – the Behavioural Assessment of the Dysexecutive Syndrome in Children (BADS-C; Emslie, Wilson, Burden, Nimmo-Smith, & Wilson, 2003), also aims to assess executive dysfunction in an ecologically valid way. However, its psychometric properties are weaker than would ideally be desired (Baron, 2007). In particular, the test-retest reliabilities of the BADS-C are highly variable and in some cases are “perhaps worryingly low” (Henry & Bettenay, 2010, p. 117). It is also worth noting that the BADS, DEX and BADS-C, like their less ecologically valid counterparts, were not designed for use in autistic populations. Therefore even the most ecologically valid of existing measures of executive dysfunction fall short.

In the context of the limitations of the measures described above, particularly in terms of their ecological validity and suitability for the ASD client group, the aims

of the current study are to develop and pilot a measure specifically designed to capture executive functioning difficulties as they present in children with ASD. It is intended that this measure:

- Is ecologically valid, having both high verisimilitude and high veridicality (Kenworthy et al., 2008);
- Has been created especially for this client group;
- Improves on the test-retest reliability of existing measures of executive functioning.

Rather than designing tasks to measure theoretical executive functioning constructs, such as ‘planning’, ‘set-shifting’ or ‘generativity’, we will take a symptom-led approach to task development. This will involve creating tests that aim to provide a standardised measure of the types of planning and organisational problems that children with ASD encounter in their daily lives. Taking this ‘bottom-up’ approach supports our objective with regard to verisimilitude and permits us not to make assumed links between observed behaviours and abstract executive function constructs. Successful development of this measure will not only be valuable in developing our understanding of the types of executive difficulties individuals with ASD typically experience (White et al., 2009), but will also be important in guiding assessment, formulation and the development of suitable interventions (Hill, 2004b).

2.3 Method

2.3.1 Participants

Ethical approval for the study was given by the Westminster Research Ethics Committee (REC reference 15/LO/1332; Appendix II) and informed consent was obtained from all participants’ parents prior to being included in the study. Children

also provided informed assent before taking part in the study. Information sheets and consent forms for parents and children can be found in Appendix III.

Forty children aged between 8 and 13 years old were recruited to take part, providing an appropriately sized sample for the purposes of this study, which aimed to initially investigate the properties of the new measure in question. Participants were divided into two groups: an ASD group ($n = 20$) and a group of typically developing age and IQ matched controls ($n = 20$). Participants were only included in the study if they were aged between 8 and 12 at the point of recruitment, had an IQ above 70, and had a good level of understanding of English. An additional criterion was applied to the ASD group: participants had to have previously received a clinical consensus diagnosis of ASD, autism, high-functioning autism (HFA) or Asperger syndrome from a qualified clinician. These diagnoses were given on the basis of information gathered using the Developmental, Dimensional and Diagnostic Interview (3Di; Skuse et al., 2004), the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989) and school reports. An additional criterion was also applied to the typically developing group: participants were excluded from the study if they had a diagnosis of a neurodevelopment disorder, such as ASD, attention deficit hyperactivity disorder (ADHD), or Tourette's syndrome. Children in the ASD group were recruited from the Social Communication Disorders Clinic at Great Ormond Street Hospital, while children in the control group were recruited predominantly through a mainstream London school. Initial contact letters for parents of control and clinical participants can be found in Appendix IV. Five children were also convenience sampled through researcher contacts. Two participants in the ASD group had comorbid diagnoses of ADHD, one of attention deficit disorder (ADD), one of generalised anxiety disorder, one of oppositional

defiant disorder, and one of obsessive compulsive disorder. None of the controls were reported to have any developmental disorders, though one child had Ehlers-Danlos Syndrome Type 3. Three of the controls had siblings with ASD, and one had a sibling with a significant speech and language disorder. Participant characteristics are summarised in Table 2.1.

2.3.2 Procedure

All of the measures listed below were initially administered by one of the two trainees conducting the research, or one of three trained research assistants, in a single 90-minute test session. Parents were given the option of completing parent questionnaires either ahead of the session or during the session. Ten clinical participants and 10 controls were randomly selected and invited to complete the Eco-TED for a second time in order to assess test-retest reliability. Participants living outside of the Greater London area were not approached to take part again because of financial and time constraints. Six clinical participants and 8 controls were available for retesting; the rest either declined or did not respond. ASD participants completed their test session at home or in a private room within the University College London campus. Controls who were recruited through their school were tested in a private room on site during the school day. Controls who were convenience sampled were tested at home. Breaks and refreshments were offered and participants received a £5 voucher as a reward for taking part, as well as entry into a raffle to win a £50 voucher.

Table 2.1

Mean (and standard deviations) for participant characteristics in both groups

	ASD group	Control group
<i>N</i> (male:female)*	20 (10:10)	20 (17:3)
Age (months)	135.90 (13.19)	131.05 (14.85)
IQ	102.06 (14.43)	107.60 (13.98)
Ethnicity	17 White British 3 Mixed Ethnicity	17 White British 2 White Other 1 Mixed Ethnicity
Clinical diagnosis	11 ASD ^a 1 HFA ^b 8 Asperger syndrome	

p = .018^a Autism Spectrum Disorder^b High Functioning Autism

Note. IQ scores for two clinical participants could not be calculated as participants did not wish to complete the necessary tests

2.3.3 Measures

2.3.3.1 Eco-TED. The central aim of this study was to develop and pilot the Ecologically-Valid Test of Executive Dysfunction (Eco-TED) for children with ASD. The subtests that comprise this measure were designed to reflect the difficulties with executive functioning that children with ASD experience in everyday life. Therefore, as mentioned above, we took a symptom-led approach in devising this measure. This involved designing Eco-TED subtests to resemble real-life situations in which children with ASD struggle with planning and organising their behaviour. To this end, task development was informed by data collected at the Social Communication Disorders Clinic using the 3Di (Skuse et al., 2004). The 3Di

assesses autistic symptomology based on parental reports and is suitable for use in both clinical and non-clinical populations. The clinic was able to provide 3Di data for a large sample of children with ASD who had attended the Social Communication Disorders Clinic in the past ten years. Data for a large non-clinical sample and a non-ASD clinical sample were also provided by the clinic. We examined the data for the 3Di items concerned with executive functioning ($n = 33$), looking at how effectively these items discriminated the ASD and non-ASD groups. In order to guide which particular behaviours to focus on in task development, we only selected those 3Di items that were significantly more endorsed for the ASD group ($n = 11$). These items are listed in Table 2.2 and formed the basis of the development of the Eco-TED subtests.

Possible approaches to measurement were thought about for each of the items listed in Table 2.2 and at least one rudimentary task proposal was developed for each. These ideas and proposals were discussed in the broader research team, which included Dr Will Mandy and Professor Paul Burgess, both of whom are experienced in the development of neuropsychological measures. The feasibility and appropriateness of each of these basic task outlines was considered carefully and a few were ruled out at this stage as a result. The remaining ideas were then developed further through an iterative process of task design and informal piloting. By the end of this process, seven subtests remained. Only three of the final subtests will be described here as this research was conducted jointly with another doctoral trainee whose thesis focuses on the remaining four subtests (see Appendix V for relevant sections of the task script; Pullinger, 2016). This paper concerns the ‘Consent Form Test’ and ‘the School Bag’ and ‘Lego Tasks’. Each trainee’s contribution to the project is detailed further in Appendix VI.

Table 2.2

Items from the 3Di that formed the basis of the development of the Eco-TED subtests

3Di item
Does X easily or frequently lose things he/she needs, for example, for school?
Is X able to tie his/her shoelaces without help?
Can X give an easy-to-follow account of past events such as a birthday party or holiday?
Can X talk clearly about what he/she plans to do in the future (e.g. tomorrow, or next week?)
Would X have difficulty in explaining to a younger child how to play a simple game?
And what about difficulty in telling a story, or describing what he/she has done, with an orderly sequence of events?
Can X remember complex commands such as ‘go upstairs, get your dirty washing, bring it down and put it in the laundry basket’?
Has X ever played a game with life-like figures or animals in which he/she talks to them?
Has X ever played a game in which there are several figures or animals and they are talking to each other?
Does X become upset by unexpected events that most children would find pleasurable?
Was there ever a time when X had to do things, or have you do things, in some precise routine?

2.3.3.1.1 Consent Form Test. This task was based on the 3Di item that enquires about the child losing things he/she needs, for example, at school. We aimed to emulate a similar situation by requiring participants to ‘remember to remember’ by recalling something presented to them as important, much like a caregiver or teacher might with the items the child misplaces. In executive functioning terms, the Consent Form Test broadly aims to assess prospective and working memory. At the beginning of the testing session the participant reads through and signs a copy of the child information and assent sheet with the help of the researcher. Once the participant has signed the assent form, it is placed in a

folder and positioned within view of the participant, on a flat surface (e.g. a table or chair) to the side of the examiner, no more than two metres away from the table on which testing is taking place. The participant is informed that they must ask the researcher for this piece of paper once their session is over, and before they leave. Approximately 90 minutes later, at the end of the session, the researcher waits five seconds to see if the participant remembers to ask for their signed assent form without prompting. If they do not, a series of prompts are given, each decreasing in ambiguity. The more specific the prompt required to remind the participant, the fewer points the participant is awarded for this task. The participant may achieve a maximum of 3 points for remembering to ask for the assent form without any prompting, and a minimum of 0 for requiring an explicit reminder.

2.3.3.1.2 School Bag Task. Like the Consent Form Test, this task was originally derived from the 3Di item that enquires about the child losing things he/she needs, for example, at school. We considered the factors that might lead the child to misplace or forget the things they need at school and aimed to recreate these demands in this task. In broad terms, the School Bag Task pertains to the executive functions of planning and prospective memory. Participants are asked to plan for activities during an imaginary day at school. An A2 sized poster is placed on the table directly in front of the participant. The poster depicts 81 everyday objects, including those that the participant would need for the given school activities, as well as distractor objects which they would not. Before commencing the task, trial (control) items are used to establish that the participant is able to recall a short list of objects and find these on the poster. In the main task, the participant is read a series of prompts outlining the activities in question. For example, Item 2 of the School Bag Task was presented to participants in the following way:

“It’s break time. A friend in your class is bringing in a cake to celebrate their birthday. They asked you to bring in some balloons for them. What do you need?”

Participants are read a maximum of seven of these prompts, which increase in complexity by increasing the number of objects the participant needs for the activity and, in some cases, by introducing a choice between two activities. Each time, the participant uses sticky labels to indicate which objects on the poster they have selected. For all items, the correct and incorrect objects selected are recorded, as well as the time taken for the participant to select their objects. For Item 2, there is only one correct item: balloons. A planning efficiency score for the entire task is calculated by subtracting the total number of incorrect items selected during the task from the total number of correct items. The total time spent on the task is also calculated. In addition, for each of the seven items a note is made of whether the participant pauses between hearing the instructions and starting to select objects: a score of 0 is awarded if the participant doesn’t pause; 1 if the participant pauses for 0-5 seconds; and 2 if the participant pauses for more than 5 seconds. These ‘time paused’ scores are totalled for each individual. This task is discontinued after 3 items are failed i.e. if the participant fails to select a single correct object from the poster within 120 seconds.

2.3.3.1.3 Lego Task. This task was based on the 3Di item that enquires about the child’s ability to remember complex commands, such as ‘go upstairs, get your dirty washing, bring it down and put it in the laundry basket’. We aimed to recreate an analogous situation, requiring participants to remember and carry out a complex set of instructions in the presence of distractors. In executive functioning terms, the Lego Task broadly aims to assess working memory and attention. In this task

participants are asked to sort through and find specifically shaped and coloured pieces of Lego. The examiner explains that they wish the participant to help them find some Lego pieces that are hidden amongst a large selection of pieces in two bowls (65 pieces in each bowl). The participant is asked to find three specific, different types of Lego pieces (a total of six pieces) as quickly as they can. Three demonstration Lego pieces are used to explain the task and the participant is given three opportunities to learn which pieces they need to find. If they are unable to recall the list perfectly after three attempts, the task is discontinued. If the participant shows that they have learnt the list of Lego pieces, they are asked to place them in a cup once they have found them. Before they commence the task, they are then told that there are also small animal figures hidden in each bowl and that they will need to select their favourite one of these for each bowl and place it by the cup containing the Lego. These animals function as distractors. The researcher then ensures that the child understands all the requirements of the task by asking them to relate their understanding of the instructions. The researcher explains any element that the child has omitted or repeated incorrectly, and then begins the task. If the child is unable to grasp what is required or fails to start the task within 60 seconds of hearing the instructions, the task is discontinued. The time taken to complete the task is recorded. In addition the contents of the cup is scored, yielding a maximum possible score of 6 and a minimum of 0, and any additional or incorrect items are totalled. Finally, the placement of the two animal figures is recorded and scored in the following way: 2 points are awarded if the participant has placed 2 animals in the correct place; 1 point if 2 animals have been selected but put in the wrong place, or if an incorrect number of animals has been placed correctly; and 0 if the participant has

not selected any animals or has put an incorrect number of animals in the wrong place.

2.3.3.2 Neuropsychological measures. In order to assess the criterion validity of the Eco-TED, existing measures of executive functioning were also included in our battery. Two subtests from the BADS-C (Emslie et al., 2003) were used for this purpose as the BADS-C arguably represents the current ‘gold standard’ in the assessment of executive functioning in children as it has been used so extensively in this field. The BADS-C has also been shown to have good discriminative and construct validity (Baron, 2007; Engel-Yeger, Josman, & Rosenblum, 2009). The Zoo Map and the Six Part subtests of the BADS-C were selected as the adult BADS (Burgess et al., 1996) equivalents of these tests have been shown to discriminate effectively between adults with ASD and controls (Hill & Bird, 2006). The Zoo Map Test involves planning and requires the participant to devise a route to visit six specified locations (out of a possible 12) on a map of a zoo whilst following a set of rules. The Six Part Test involves planning, task-scheduling and monitoring of action: participants have five minutes to carry out a set of tasks and must follow a set of rules while they do this.

In addition, participant IQ was assessed using the two-subtest version of the Wechsler Abbreviated Scale of Intelligence – Second Edition (WASI-II; Wechsler, 2011). The WASI-II has evidenced reliability and validity (Irby & Floyd, 2013), and its two subtest version correlates strongly with the Wechsler Intelligence Scale for Children-Fourth Edition full scale IQ ($r = 0.83$; Homack & Reynolds, 2007). Five of our clinical participants had already gone through IQ assessments as part of their assessment with the Social Communication Disorders Clinic. In these cases, existing scores were used.

2.3.3.3 Parent-report questionnaires.

2.3.3.3.1 *The Behaviour Rating Inventory of Executive Functions (BRIEF;* Gioia, Isquith, Guy, & Kenworthy, 2000). Parents completed the BRIEF - Parent form (Gioia et al., 2000). This is a standardised parent- or teacher-report rating scale and is composed of 86 items in which the parent responds whether their child has difficulties with specific behaviours: Never, Sometimes, or Often, scored as 1, 2, or 3, respectively (Gioia et al., 2000). Higher scores indicate greater symptom severity. The BRIEF was developed to assess behavioural manifestations of children's executive control functions in their everyday activities. It was used here as an additional means of assessing the criterion validity of the Eco-TED. In addition, as it is an informant report measure of everyday activities, it also provided a way of assessing the veridicality of our tasks.

2.3.3.3.2 *Social and Communication Disorders Checklist (SCDC; Skuse et al., 1997).* Participants' parents also completed the SCDC: a brief, reliable and valid screening questionnaire for autistic traits (Skuse, Mandy, & Scourfield, 2005). It comprises 12 questions on the social and communication aspects of ASD, and has excellent sensitivity (.88) and specificity (.91) when discriminating ASD from non-ASD cases (Skuse et al., 2009). In addition, the SCDC is a very reliable instrument, with a high internal consistency ($\alpha = .93$) and test-retest reliability ($r = .81$ over two years)(Skuse et al., 2005). It was used here to measure autistic traits among our clinical participants and to screen controls for possible social communication difficulties.

2.3.3.3.3 *Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997).* Parents of all participants completed the SDQ (Appendix VII; Goodman, 1997). This is a brief behavioural screening questionnaire for 3-16 year olds. It has been

shown to have good predictive validity across a range of clinical disorders (Goodman & Goodman, 2009). The SDQ enquires about: emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems; and prosocial behaviour. Items pertaining to the first four domains are added together to generate a total difficulties score. Higher scores indicate greater symptom severity. The SDQ also includes an optional additional subscale that aims to assess the impact of reported difficulties on the child's life. In this study, the SDQ served as a means of characterising our sample and therefore assessing the generalisability of our findings.

2.3.4. Data analysis

Statistical analysis was conducted using SPSS version 22.0. Before running parametric tests heterogeneity and normality were checked. Where these assumptions were not met, non-parametric alternatives were used. Our analyses are detailed below.

2.3.4.1 Data analysis procedures.

2.3.4.1.1 Group differences. Group differences for SDQ Overall Stress and SCDC Total scores were explored using Mann Whitney *U* tests. In addition, non-parametric analyses were used to assess group differences for our measures of executive functioning: the BRIEF, the Zoo Map and Six Part subtests of the BADS-C, and the Eco-TED. A variety of different scores were analysed for the Eco-TED (see below). For the established measures of executive functioning, we used composite scores where possible (BRIEF) and scaled scores when composite scores could not be calculated (BADS-C).

2.3.4.1.1.1 Eco-TED raw scores. In order to establish whether the Eco-TED subtests could distinguish between our ASD and non-ASD groups on the basis of raw scores, we conducted non-parametric group-level analyses using Consent Form

Test raw scores and raw composite scores for the School Bag and Lego Tasks (n = 8).

2.3.4.1.1.2 *Eco-TED deficit scores.* The successful development of neuropsychological measures involves not only task development but also consideration of the appropriate means of scoring the tasks and analysing the data they yield. Additionally, for the data yielded to be informative, it must be compared to the performance of an appropriate normative sample (Weiner, 2003). As Anastasi & Urbina (1997, p. 67) state, “in the absence of additional interpretative data, a raw score on any psychological test is meaningless.” As well as conducting analyses using Eco-TED raw scores, we therefore considered different empirical approaches and ultimately sought to compare participant scores to norms based on the performance of the current control group. We used a method very similar to that described by Towgood et al. (2009). Based on the neuropsychological single-case design which has been widely used in the field of cognitive neuropsychology, Towgood et al.'s (2009) approach similarly involves the comparison of individual performance with norms derived from a control group.

The central aim for this stage of the analysis was to define ‘normal’ and ‘abnormal’ performance for each Eco-TED task, based on the performance of the control group, and then compare each individual’s performance to this normative standard. We began this process by examining the raw Consent Form Test scores and raw composite scores for the School Bag and Lego Tasks and correcting any outliers by winsorizing them. When winsorizing, outliers can be corrected in a number of different ways. In this instance, the simplest approach was taken and outliers were replaced with the next highest/lowest score in the sample that was not an outlier. Three data points were corrected by winsorizing. Subsequently, the

distributions of the data were examined again using statistical tests. After winsorizing, five of the eight Eco-TED outcome measures fell within the acceptable ranges for skewness, and two fell within the acceptable ranges for kurtosis.

In order to conduct the next stage of the analysis, raw Consent Form Test scores and raw composite scores for the School Bag and Lego Tasks were converted to z -scores. These z -scores were based on the mean performance of our IQ and age matched control group, thus providing scores for all participants relative to “normal” cognitive performance. As in Towgood et al.'s (2009) paper, these z -scores were then used for each of the eight Eco-TED measures to identify when participants showed abnormal performance, defined as two standard deviations below the control group mean. These individuals were awarded a score according to how many standard deviations below this cut off they performed. This score represented the individual's “deficit” for that particular measure and served as a numerical identifier for those participants who performed outside of the normal range. Those who performed within the normal range were awarded a score of 0. As well as the deficit scores that were calculated as above, ‘total deficit scores’ for the Lego and School Bag Tasks were calculated for each participant. This was done by totalling each individual's full complement of deficit scores for the particular task. For example, a total deficit score for the School Bag Task was calculated for each participant by adding their deficit scores for planning efficiency, total time taken and total time paused. This could not be done for the Consent Form Test as it provided only one outcome measure.

Mann Whitney U tests were applied to those deficit scores where at least some participants showed a deficit ($n = 9$). None of our participants scored in the impaired range (two standard deviations below the control group mean) for the

Consent Form Test meaning that they were all awarded a deficit score of 0. The Consent Form Test deficit scores were therefore excluded from the analysis at this stage.

2.3.4.1.1.3 School Bag Task individual items. As part of our exploration of the properties of the Eco-TED we were also interested in seeing which of the School Bag Task individual items were most effective in terms of discriminating between our groups. This would provide useful information on which to base any further development of this task. Group differences between control and ASD raw scores were analysed using Mann Whitney *U* tests for a total of 28 measures derived from the School Bag Task. These included the four possible sub-measures (correct items selected, additional or incorrect items selected, total time taken and the time paused before beginning to select items) for each of the seven School Bag Task items. Conducting this large number of comparisons increased the likelihood of making a type I error. This is discussed below with reference to our findings.

2.3.4.1.1.4 Individual variation in performance. As well as looking at how each individual's performance compared with our control group, we also wished to explore within-person variability across tasks. As mentioned above, a tendency towards having a pattern of peaks and troughs in performance is considered by some as a hallmark of ASD (e.g. Nydén et al., 2011). In fact, Towgood et al. (2009) found variability in performance to be the most defining feature of their ASD group when they explored how individual performance varied across tasks. In a similar way, we sought here to explore individuals' profile of abilities and see how these compared to our control group in terms of their variability. Mann Whitney *U* tests were used for this purpose. Individual variability could not be assessed for the Consent Form Test as there was only one outcome measure for this test. Our hope was that the Eco-

TED tasks would capture within-person heterogeneity in the ASD group. This would provide evidence for the appropriateness of the use of the Eco-TED with this population

2.3.4.1.2 Correlations. Non-parametric correlations were used to explore associations (by group) between Eco-TED raw scores and deficit scores and SDQ Overall Stress and SCDC Total scores. Non-parametric analyses were also used to assess correlations with participant age and IQ across both groups, as well as with BRIEF GEC scores and BADS-C subtest scores for each group. The latter served as a way of assessing the criterion validity of the Eco-TED tasks and our empirically derived method of scoring. Finally, test-retest reliability was evaluated using non-parametric correlations between initial and retest raw scores.

2.4 Results

2.4.1. Group differences

2.4.1.1 SDQ and SCDC. Both the SCDC Total and SDQ Overall Stress scores were significantly higher in the ASD group (see Table 2.3).

2.4.1.2 Executive functioning.

2.4.1.2.1 BRIEF. BRIEF GEC scores were significantly higher among ASD participants (see Table 2.3).

2.4.1.2.2 BADS-C. Although the control group scored higher on average than ASD participants on the BADS-C Zoo Map and Six Part Tests, this difference was not significant (see Table 2.3).

2.4.1.2.3 Eco-TED.

2.4.1.2.3.1 Eco-TED raw scores. As can be seen from Table 2.4, no significant group differences were found for any of the eight Eco-TED raw measures. However, there were non-significant trends for the control group to

perform better on average than the ASD group on all but one measure: the number of correct pieces selected in the Lego task.

2.4.1.2.3.2 Eco-TED deficit scores. School Bag Task total deficit scores were significantly higher in the ASD group, $U = 310, p = .002$. School Bag Task total time paused deficit scores were also significantly higher among ASD participants, $U = 280, p = .030$. A total deficit score excluding participants' total time paused deficit scores was also calculated for each participant. These scores were higher in the ASD group ($M = 1.20$) than the control group ($M = 0.00$) but this difference was not significant, $U = 250, p = .183$.

With regard to the Lego Task, although single composite deficit scores were not significantly different between our groups, total deficit scores were higher in the ASD group, $U = 306, p = .004$. A total deficit score excluding the time taken deficit scores was calculated for each participant. Although these scores were lower among controls ($M = 0.14$) than ASD participants ($M = 0.61$), this difference was not significant, $U = 265, p = .081$. These results are summarised in Table 2.4.

2.4.1.2.3.3 School Bag Task individual items. Only the time taken to complete Item 6 of the School Bag Task showed a significant group difference, $U = 313, p < .001$, with ASD participants taking longer ($M = 45.21$) than controls ($M = 24.60$). The number of comparisons carried out increased the risk of making a type I error at this stage. However, as our only finding was very significant, it is unlikely that the observed effect was a 'false positive'. These results are summarised in Appendix VIII.

2.4.1.2.3.4 Individual variation in performance. Variability in performance was initially explored using line graphs of the z-scores calculated as above for both groups. As can be seen in Figures 2.1 and 2.2, within-person variation in

performance across the various Eco-TED measures and tasks appeared greater for ASD participants. For the School Bag Task, individual variability between items and sub-measures was significantly higher among ASD participants ($M = 7.63$) than controls ($M = 3.26$), $U = 328.50$, $p < .001$. The highest and lowest z-scores for this sample were represented in the School Bag Task results so the individual variability across all three Eco-TED tasks was equal to that for this task. For the Lego Task, individual variability between sub-measures was significantly less among controls ($M = 1.80$) than ASD individuals ($M = 4.37$), $U = 304$, $p = .004$.

Table 2.3

Group differences, means, standard deviations, p values and effect sizes for the SCDC, SDQ, Zoo Map and Six Part Tests, and the BRIEF

Measure	N	Control mean (sd)	ASD mean (sd)	p	Effect size (d)
SCDC Total	40	1.50 (1.85)	16.65 (4.84)	.000	4.13
SDQ Overall Stress	40	4.65 (4.27)	18.80 (6.86)	.000	2.48
BADS Zoo Map 1	40	8.45 (3.12)	8.40 (4.03)	.718	0.01
BADS Zoo Map 2	36	10.44 (3.25)	7.90 (4.36)	.072	0.66
BADS Six Part Test	40	9.45 (3.28)	8.95 (3.46)	.964	0.15
BRIEF Global	40	44.40 (6.48)	74.05 (9.89)	.000	3.55
Executive Composite					

Table 2.4

Group differences, means, standard deviations, *p* values and effect sizes for the raw and deficit Eco-TED scores

Task	Raw or deficit	Measure	<i>N</i>	Control mean (sd)	ASD mean (sd)	<i>p</i>	Effect size (<i>d</i>)
Consent Form	Raw	-	38	1.21 (0.92)	1 (1.00)	.435	-0.22
School Bag	Raw	Planning efficiency	40	7.89 (6.25)	4.90 (11.75)	.565	-0.32
School Bag	Raw	Total time taken	40	194.93 (65.42)	265.00 (135.85)	.142	0.66
School Bag	Raw	Total time paused	40	4.85 (2.32)	3.60 (4.06)	.108	-0.38
Lego	Raw	Total time	40	36.19 (10.80)	53.30 (40.25)	.063	0.58
Lego	Raw	Correct pieces	40	5.60 (1.05)	5.60 (0.88)	.841	0
Lego	Raw	Incorrect pieces	40	0.35 (0.99)	0.45 (0.89)	.620	0.11
Lego	Raw	Animal placement	40	1.90 (0.31)	1.60 (0.60)	.174	-0.63
School Bag	Deficit	Planning efficiency	40	0.00 (0.00)	0.68 (1.95)	.289	-0.50
School Bag	Deficit	Total time taken	40	0.00 (0.00)	0.52 (1.25)	.289	-0.59
School Bag	Deficit	Total time paused	40	0.00 (0.00)	0.04 (0.05)	.030	-1.12
School Bag	Deficit	Total deficit:	40	0.00	1.24	.002	-0.57

				Planning efficiency,	(0.00)	(3.07)	
				total time taken and			
				total time paused			
Lego	Deficit	Total time	40	0.00	1.11	.060	-0.57
				(0.00)	(2.76)		
Lego	Deficit	Correct pieces	40	0.05	0.12	.968	-0.24
				(0.12)	(0.38)		
Lego	Deficit	Incorrect pieces	40	0.00	0.00	1.000	-
				(0.00)	(0.00)		
Lego	Deficit	Animal placement	40	0.09	0.49	.175	-0.55
				(0.29)	(0.97)		
Lego	Deficit	Total deficit: Total	40	0.14	1.72	.004	-0.78
		time, correct pieces,		(0.35)	(2.84)		
		incorrect pieces and					
		animal placement					

Figure 2.1

Control participants' profile of performance across the 34 Eco-TED sub-measures

for which z-scores could be calculated

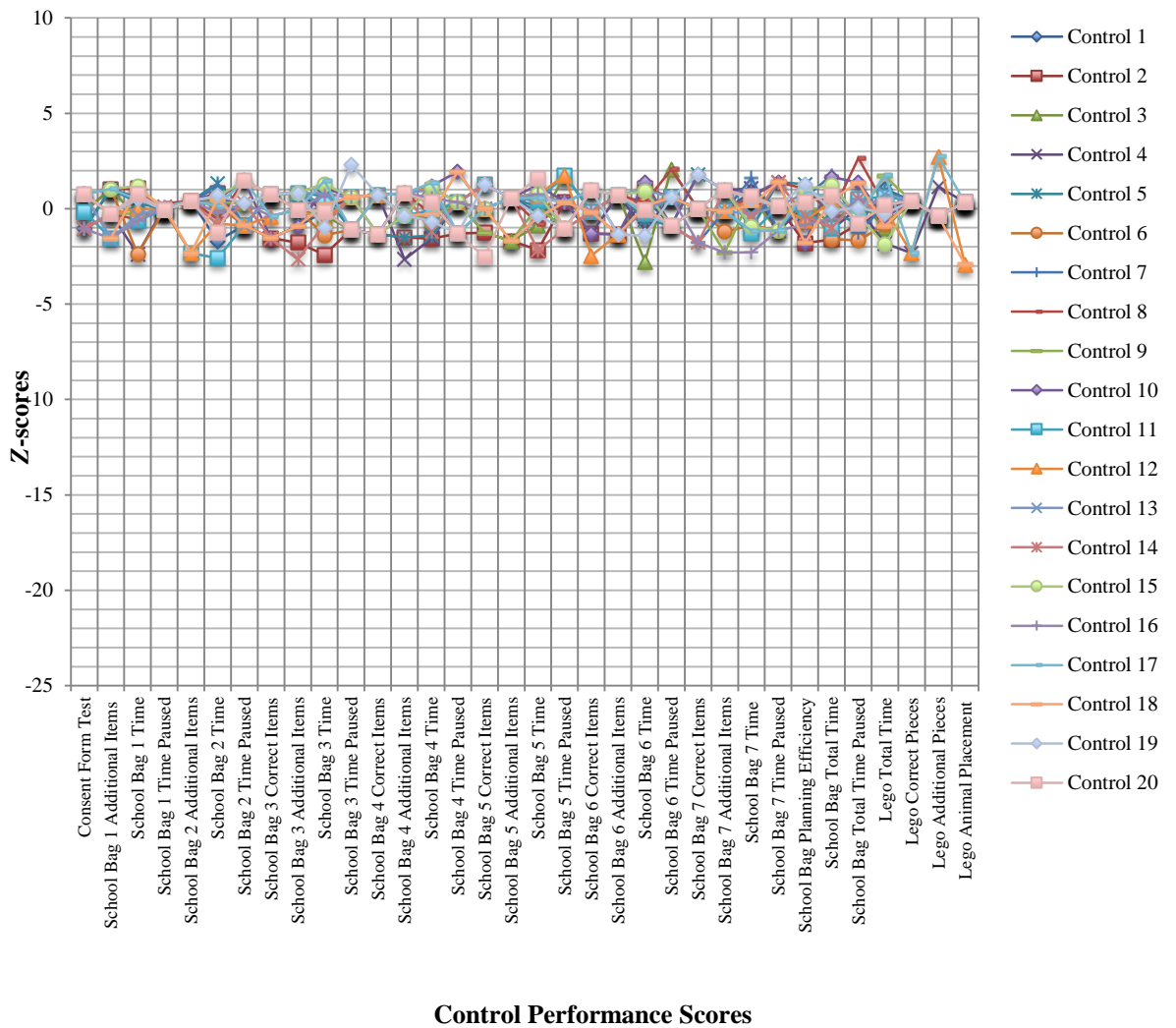
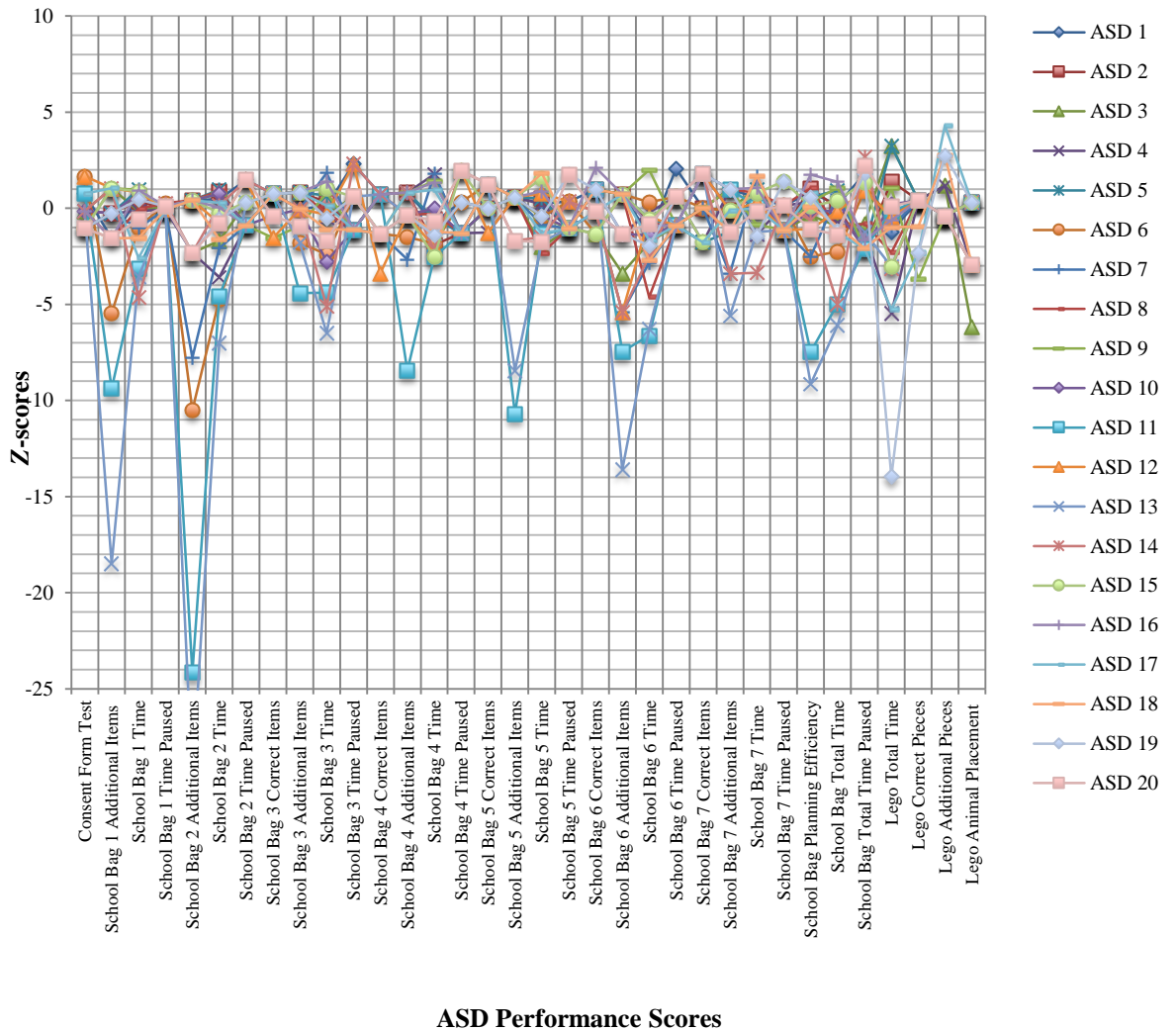


Figure 2.2

ASD participants' profile of performance across the 34 Eco-TED sub-measures for which z-scores could be calculated



2.4.2 Correlations

It is important to note that the multiple correlations carried out increased the risk of type I error. However, Bonferroni corrections were not applied to our p values as they can be too conservative and therefore increase the risk of type II errors (Perneger, 1998).

2.4.2.1 Correlations with SDQ and SCDC. The SDQ Overall Stress and SCDC Total scores correlated significantly with control group scores for the Consent Form Test. SCDC scores also correlated with ASD raw scores for the Lego Task correct pieces measure. These results are summarised in Table 2.5. There were no other significant results.

2.4.2.2 Correlations with IQ and age. Participant age correlated significantly with the School Bag Task planning efficiency raw composite measure, $r_s = .327$, $p = .039$. In addition, age correlated with the Lego Task total time taken deficit composite, $r_s = .344$ $p = .030$. IQ scores did not correlate significantly with any of our Eco-TED measures.

2.4.2.3 Correlations with established measures of executive functioning. The BRIEF GEC correlated significantly with control group scores for the Consent Form Test. Zoo Map Test 2 scores also correlated significantly with ASD raw scores for the School Bag Task total time paused composite. In addition, ASD group scores for the Six Part Test were found to correlate significantly with the raw Lego Task total time measure. These results are summarised in Table 2.5. It was not possible to run analyses for many of the deficit measures, particularly for the control group, as the vast majority scored the minimum of 0. This was the case for control group scores for all of the School Bag Task deficit measures (planning efficiency, total time taken, total time paused and total deficit), and for the Lego Task total time taken and

additional or incorrect pieces measures. It was also an issue for ASD deficit scores for the Lego Task additional or incorrect pieces measure.

2.4.2.4 Test-retest reliability. None of the test-retest correlations of the Eco-TED reached the minimum of .80 required for a clinical psychometric instrument (Anastasi & Urbina, 1997). The correlations varied between .013 and .270 (Table 2.6). It was not possible to run analyses for the Lego Task animal placement measure as there was insufficient variation among participant scores.

2.4.2.5 Post hoc correlations. Non-parametric correlations were used to explore associations between total time paused deficit scores for the School Bag Task and hyperactivity (as measured by the SDQ hyperactivity subscale) in our ASD group. These scores did not correlate, $r_s = .281$, $p = .078$. In addition, non-parametric analyses were used to examine the relationship between total time paused deficit scores for the School Bag Task and performance in the ASD group on the other School Bag Task measures. Total time paused deficit scores did not correlate with planning efficiency ($r_s = -.117$, $p = .623$), or total time taken ($r_s = .324$, $p = .164$) deficit scores. The rationale for conducting these analyses is outlined in the discussion below.

Table 2.5

Correlation coefficients and significance values by group for the Eco-TED measures and SDQ Overall Stress, SCDC Total, Zoo Map and Six Part Tests, and BRIEF Global Executive Composite scores

Group	Task	Raw orMeasure	SDQ	SCDC	Zoo Map 1	Zoo Map 2	Six Part	BRIEF
Controls	Consent Form	Raw -	-.487*	-.564*	.166	.488	.116	-.457*
	School Bag	Raw Planning efficiency	-.268	.010	.052	.008	-.242	-.212
	School Bag	Raw Total time taken	.134	.315	-.062	-.186	.155	.416
	School Bag	Raw Total time paused	-.035	.049	.241	.328	-.145	.096
	Lego	Raw Total time	.181	-.201	-.247	.209	.170	.311
	Lego	Raw Correct pieces	-.131	.155	-.096	-.175	-.228	-.035
	Lego	Raw Incorrect pieces	.129	-.170	.119	.175	.227	.027
	Lego	Raw Animal placement	-.131	.361	-.190	-.395	-.088	.159
	Lego	Deficit Correct pieces	.172	-.139	.086	.176	.258	.073
	Lego	Deficit Animal placement	.131	-.361	.190	.395	.088	-.159
	Lego	Deficit Total deficit: time, correct pieces, incorrect pieces and animal placement	Total.118	-.286	.061	.407	.144	-.118
ASD	Consent Form	Raw -	.026	.243	.279	.157	-.246	.301
	School Bag	Raw Planning efficiency	-.186	-.160	.029	.139	-.060	-.091
	School Bag	Raw Total time taken	.065	-.187	-.116	-.098	-.371	.063
	School Bag	Raw Total time paused	.400	.151	-.242	-.487*	-.236	.271
	School Bag	Deficit Planning efficiency	-.186	.033	.134	.094	-.109	-.056

School Bag Deficit	Total time taken	.002	-.092	.242	.149	-.384	.090	
School Bag Deficit	Total time paused	-.302	-.080	.151	.404	.304	-.293	
School Bag Deficit	Total deficit: Planning efficiency, total time taken and total time paused	-.120	-.096	.283	.332	-.204	-.018	
Lego	Raw	Total time	-.135	-.332	-.174	-.256	-.702**	.102
Lego	Raw	Correct pieces	.227	.471*	.431	.033	.113	.261
Lego	Raw	Incorrect pieces	.000	-.230	-.360	.051	-.187	.095
Lego	Raw	Animal placement	-.059	.201	.397	.289	-.342	.025
Lego	Deficit	Total time	-.085	-.300	-.048	-.090	-.682	.055
Lego	Deficit	Correct pieces	-.203	-.337	-.323	.066	-.100	-.251
Lego	Deficit	Incorrect pieces						
Lego	Deficit	Animal placement	.059	-.201	-.397	-.289	.342	-.025
Lego	Deficit	Total deficit: Total time, correct pieces, incorrect pieces and animal placement	-.022	-.447	-.426	-.142	-.260	-.043

* $p < .05$

** $p < .001$

Table 2.6

Correlation coefficients for the initial Eco-TED raw scores and retest Eco-TED raw scores

Task	Measure	Retest scores
Consent Form	-	.013
School Bag	Planning efficiency	-.264
School Bag	Total time taken	.270
School Bag	Total time paused	-.193
Lego	Total time	.176
Lego	Correct pieces	-.077
Lego	Incorrect pieces	-.166

2.5 Discussion

This paper describes the development and subsequent pilot of the Eco-TED; a new measure of executive functioning designed especially for children with ASD. Twenty children with ASD and 20 age and IQ matched controls completed the Eco-TED. The results of this pilot will be discussed with reference to the three Eco-TED tests this paper is concerned with: the Consent Form Test and the School Bag and Lego Tasks. Their strengths and limitations will be considered, as well as any recommendations for future research. The study's limitations and strengths more broadly will then also be considered, and finally conclusions on the study as a whole set out.

2.5.1. The Eco-TED pilot: Summary and interpretation of results

Group-level analyses of raw Eco-TED scores revealed little in the way of group differences. However, when we applied an empirical method that permitted us

to consider participants' performance relative to the control group, some important group differences were identified. Following Towgood et al.'s (2009) example, we based our analysis on performance scores which were calculated relative to the performance of our control sample. By using our control group scores to set the standard for both sub-normal and normal performance we were able to calculate each individual's "deficit". We found that the ASD group scored more highly in terms of deficit or impaired performance on two of the Eco-TED tasks. These findings are discussed below in more detail but it is important to note that although there is a precedent for using this kind of method in the field of cognitive neuropsychology, these analyses are exploratory and our findings would therefore need to be replicated in an independent sample. In addition, we explored correlations between Eco-TED composite scores and the standardised measures we administered as well as participant age and IQ.

2.5.1.1 Consent Form Test. The raw scores for the Consent Form Test indicated that both our clinical and control groups found it equally difficult. Although there was some variability in performance on this task, there was little variation between our two groups meaning that analyses using deficit scores could not be conducted. Interestingly, only two participants achieved the maximum score for this test by spontaneously remembering to ask the researcher for their signed consent form at the end of their testing session. Both of these individuals had ASD diagnoses. It would be interesting to explore whether supra-normal performance on the Consent Form Test can be observed in other individuals with ASD, and in particular whether this would form part of a mixed profile of abilities across various domains as might be expected among individuals with ASD (Towgood et al., 2009). It appears that overall this test presented too much of a challenge for our participants.

It could be that any group differences were masked by a floor effect and it would be interesting to devise an easier version of this task to see whether group differences might then emerge.

Referring back to Kenworthy et al.'s (2008) framework for assessing ecological validity, the Consent Form Test appears to have some veridicality: controls' raw scores correlated negatively with BRIEF GEC, SDQ Overall Stress and SCDC Total scores. If typically developing individuals who struggled with the Consent Form Test also had executive, social communication-type and behavioural difficulties as reported by their parents, it seems likely that the Consent Form Test does in some way capture executive-type or at least ASD-related difficulties although we have been unable to do so with our clinical participants here. Consent Form Test scores did not correlate with performance on the BADS-C Zoo Map or Six Part Tests, but we did not find any group differences for these tests suggesting that they did not capture executive functioning difficulties very well in our sample. Our clinical and control groups achieved similar means for the Zoo Map 1 and Six Part Tests, though scores for the Zoo Map 2 Test were higher among controls. This may mean that these scores do not necessarily provide a good 'criterion' against which to assess the validity of the Eco-TED. Our results may relate partly to issues with administration of these tests, particularly in relation to the Zoo Map 1 and Six Part Tests which are arguably the more challenging of the three to administer as they require greater examiner involvement. The author of this paper tested the majority of the control group, while most of the ASD participants were tested by research assistants. This might account for these results at least in part and it would be important to assess interrater reliability in future piloting of the Eco-TED.

2.5.1.2 School Bag Task. Analysis of deficit scores for the School Bag Task revealed that the task as a whole was able to distinguish between children with ASD and typically developing controls in an important way. We found that ASD participants had a greater deficit with regard to the total time they paused over the course of the task. This means that when all seven items were considered together, ASD children paused for less time than controls between hearing the instructions and starting to select the items they needed for the school day. When we combined all three composite deficit scores for the School Bag Task (planning efficiency, total time taken and total time paused) to calculate a total deficit score we found that our ASD participants were more impaired than our control group. However, this difference was no longer significant when we removed the total time paused deficit scores from our calculations suggesting that this was driving the group difference on the total deficit score.

The School Bag Task was therefore able to detect a particular feature of our ASD sample: they proceeded with the task of selecting items from the poster faster relative to controls. If we consider this behaviour in terms executive functioning and the demands of the task, it seems likely that inhibition is implicated. Taking time to think, plan, or scan the poster once the prompt was delivered would arguably have formed part of a goal-directed approach to this task where participants are told specifically that they must only pack exactly what they will need for each activity. Pausing for less time, our ASD participants appear to have been less able to suppress their responses to stimuli in order to help them achieve their goal (Verbruggen & Logan, 2008). The size of our reported effect was large in comparison to others in the literature on inhibition in ASD (e.g. $d = -1.12$ vs. $d < 0.8$; Sinzig, Morsch, Bruning, Schmidt, & Lehmkuhl, 2008). It is worth noting that although inhibition is

also implicated in ADHD and ADD (e.g. Sergeant et al., 2002) which are often comorbid with ASD (Leyfer et al., 2006), only three of our ASD participants had these diagnoses so it is unlikely that this significantly biased our findings. In addition, scores for the hyperactivity subscale of the SDQ did not correlate with this sub-measure suggesting that in any case hyperactivity was not associated with poor behavioural inhibition in the School Bag Task. Nonetheless, our findings would need to be replicated with a larger sample, using a more precise measure of the time individuals paused before beginning to select items. Although pausing before starting to select items did not correlate with either planning efficiency scores or total time taken in our ASD group, more exact measurements would allow for further exploration of the ways in which inhibition might be associated with successful execution in the School Bag Task. This might be clinically useful; for example if low planning efficiency in the School Bag Task were found to be associated with poor inhibition, one might propose interventions that target both inhibition and planning to help those individuals that struggle in one or both of these domains.

A further finding of interest was that children in our ASD group took significantly longer than controls to complete Item 6 of the School Bag Task. This was the only group difference that was found for the School Bag Task individual item scores, suggesting that this item was somehow particularly potent. Item 6 was presented to children as follows:

“It’s P.E. You can choose between tennis and football. You’ll need shoes, socks, shorts and a top: white for tennis, red for football. School provides everything else. What do you need to bring to school?”

Item 6 required participants to not only choose between two alternative activities at school, but also to ignore seemingly appropriate distractor items that were pictured

on the poster: in this case a tennis racquet, a football, tennis balls and shin pads. Other School Bag Task Items presented children with one of these demands, and Item 7 involved both, but only Item 6 required participants to focus on one of two very similar sets of items i.e. shoes, socks, shorts and a top. This item's power as a discriminator between ASD and typically developing controls merits further investigation and at the very least Item 6 could provide a template for the development of additional School Bag Task items.

In addition to differences in deficit scores, we also explored how individual variation in performance compared between groups. As we had predicted, the School Bag Task captured the pattern of marked variability we expected in our ASD group. It is possible that, rather than reflecting a variable profile of abilities, these results relate to increasing inattention in our ASD group over the course of the testing session (Leitner, 2014). However, the pattern we observed (Figure 2.2) seems to suggest a trend of peaks and troughs rather than a gradual decline as one might expect over time. These findings would need to be replicated but this feature may make the task particularly relevant to current neuropsychological research which is increasingly focused on heterogeneity within as well as between individuals (Willcutt, Sonuga-Barke, Nigg, & Sergeant, 2008). Additionally, this significant individual variation would suggest that School Bag Task performance is not simply a product of some other characteristic of our participants, such as their IQ or age. Indeed, only one of the seven possible School Bag task composite measures correlated with participant age, and none correlated with IQ.

BRIEF GEC scores did not correlate with either raw or deficit scores for the School Bag Task. Additionally, School Bag Task scores did not on the whole correlate with scores from the BADS-C subtests, though ASD raw scores for the total

time paused measure correlated significantly with scores for the Zoo Map 2 test. With regard to the empirically derived scoring method we have used here, our results suggest that this may have low validity. It was however difficult to assess this comprehensively as the large number of individuals that scored the minimum of 0, particularly among our controls, made it impossible to run correlational analyses. As mentioned above, variability in administration might account for unexpected results for the BADS-C, but the results with regard to the BRIEF are concerning and suggest that the School Bag Task's veridicality would need to be a focus in further research.

2.5.1.3 Lego Task. We found that ASD individuals' total deficit scores for the Lego Task were higher on average than controls'. This was a large effect ($d = -0.78$) and comparable to the average effect size ($d = 0.98$) reported by Pennington & Ozonoff (1996) in their review of 14 studies investigating a variety of executive functions in children with ASD. The total deficit score combined participants' total time, the number of correct and incorrect Lego pieces they selected, and their ability to find and place animal figures as instructed. Although we did not find a group effect for any single sub-measure of the Lego Task, it seems as though the time taken made the most important contribution to this effect as when we removed this sub-measure from our calculations, total deficit scores were no longer significantly higher among our ASD participants. Although time taken to complete the Lego Task seems to have been important, there seems to have been a cumulative "deficit" effect across all four Lego Task sub-measures. This is interesting and it does seem as though there is significant overlap between the sub-measures in terms of the executive domains they may tap. For example, an individual's ability to remember and select the right Lego pieces relates to their working memory but not getting distracted when selecting animals requires cognitive control. Similarly, when

selecting the animals and placing them, individuals' working memory allows them to recall the instructions about placement, and interference control to not get distracted by the surplus animals. It is worth noting that processing speed is a significant area of weakness for children with ASD (Oliveras-Rentas, Kenworthy, Roberson, Martin, & Wallace, 2012) and it is therefore possible that this contributed to the effects reported here, particularly in light of the seemingly important role of time. It would be interesting to control for this by including a test of speed of processing, such as a digit-symbol coding exercise, in future research involving the Eco-TED. It would also be useful to further investigate which of the Lego Task sub-measures cluster together using principle components analysis with a larger sample.

As with the School Bag Task, we explored how individual variation in performance on the Lego Task compared between groups. The Lego Task also successfully captured the variability we expected to see among our ASD participants, making it a more relevant and appropriate measure for this population. Only one of the nine Lego Task composite scores we analysed correlated with participant age, and none correlated with IQ, suggesting that like the School Bag Task, it successfully measures something other than participants' general ability or age.

It was concerning to find that Lego Task raw and deficit scores did not correlate with BRIEF GEC scores, and further development of the task would need to address this. The raw scores for the number of correct pieces selected by the ASD group did however correlate significantly with scores for the SCDC, suggesting that the Lego Task may measure some social-communication difficulties. Scores for the Lego Task did not generally correlate with BADS-C subtest scores. Issues regarding the administration of these tests that might account for this are discussed above, but

it is likely that the veridicality of the Lego Task and validity of our deficit scoring method need improvement and further investigation.

2.5.1.4 Test-retest reliability. Test-retest correlation coefficients for the Eco-TED raw scores were low and none reached the minimum of .80 suggested by Anastasi & Urbina (1997). This is concerning as one of our intentions in developing the Eco-TED was to improve on the test-retest reliability of existing measures of executive functioning. Test-retest reliability is an important psychometric property as it has implications for the validity of test results and therefore our ability to interpret them with confidence. However, it may be that achieving high test-retest correlations was made difficult by the very nature of the Eco-TED tasks. One of the key requirements of ecologically valid tests is that they present a novel challenge as it is “with new situations or challenges that patients with (executive dysfunction) are likely to have most difficulty” (Emslie et al., 2003, p. 22). As the Eco-TED tasks were no longer new to our participants in their second testing session, it follows that they performed differently and therefore that test-retest reliability was affected. It appears that to an extent ecological validity may come at a cost to test-retest reliability and that the Eco-TED tasks were therefore unable to overcome this obstacle. As the Eco-TED is a very new measure and in the early stages of its development, test results should nevertheless be interpreted with caution. If high test-retest reliability is unachievable in the realm of ecologically valid tests, it may be that more concerted effort needs to go into controlling for or limiting the role of potentially confounding variables (e.g. participant effort) in future piloting of the Eco-TED.

2.5.2 Limitations and strengths of the study

2.5.2.1 Limitations. A key limitation of this study was our sample size. Firstly, our sample size was smaller than has typically been seen in previous research in this area. For example, in a study which explored the ecological, concurrent and construct validity of the BADS, Norris & Tate (2000) had a sample of 73 participants (Norris & Tate, 2000). In addition, the sample size for this study would ideally have been based on Cohen's (1992) guidance on power calculations for statistical analysis involving correlation coefficients as much of our key analyses and non-significant findings involved these. When the expected effect size is large and α is set at .01, a sample size of 41 in each group is usually required to achieve sufficient statistical power (>80%). As a substantial amount of time was spent developing the Eco-TED tasks and funding for the project was limited, it was not possible to conduct the current study on a larger scale. However, doing so in future might help to build on our understanding of the Eco-TED's veridicality and improve the Eco-TED's test-retest reliability.

Certain features of our sample also merit consideration. In order to reduce the potentially confounding influence of intellectual ability we only included participants with IQs within the normal range. This affects the representativeness of our sample and therefore the generalisability of our findings as intellectual disability affects 25-40% of individuals with ASD (Baird et al., 2000; Baird et al., 2006; Chakrabarti & Fombonne, 2001). In addition, unusually, our ASD group had an even split of male to female participants. In contrast, our control group was almost entirely male. This may have biased our findings as there are behavioural sex differences between males and females with ASD. For example, females with ASD have superficially better socio-communication skills (Lai et al., 2011) which might

have advantaged our female participants with ASD and thus biased our findings. The ratio of girls to boys in our ASD group may also affect the generalisability of our findings as ASD is three times more likely to affect males than females (Baird et al., 2006). However, it is worth noting that children in the ASD group had higher SCDC and SDQ scores suggesting that our sample was at least representative of the broader population in terms of their social communication and behavioural difficulties.

As mentioned in the discussion of the results above, the division of testing between the author and research assistants may have had some effect on our findings. It would be important to give this more careful consideration in future research as it has impacted on our ability to draw conclusions about the construct validity of the Eco-TED, and may also have some implications for our interpretation of the pilot results more broadly.

As individuals with ASD have evidenced difficulties with socially administered tasks relative to computerised tasks (S Ozonoff, 1995), one might argue that the format of the Eco-TED is not ideal for this client group. However, impaired performance in our ASD group was by no means universal both between and within participants, suggesting that the way the tasks were presented did not lead to impaired performance across the board. Computerised tasks bearing little resemblance to the challenging situations we wished to recreate would also have had obvious implications for the ecological validity of the Eco-TED.

2.5.2.2 Strengths. The greatest strength of this study is the way in which the Eco-TED was devised. With Kenworthy et al.'s (2008) framework in mind, the tasks' verisimilitude, or their resemblance to real-life demands, was prioritised and the chances of designing tasks with ecological validity therefore arguably increased.

There is widespread agreement that the ecological validity of measures of executive functioning is paramount (Kenworthy et al., 2008) and that it relates directly to their clinical usefulness (Burgess et al., 2006). The Eco-TED is therefore valuable in that it is the first ecologically valid test of executive functioning that has been devised for children with ASD.

The School Bag and Lego Tasks yield a number of different types of measure as well as composite scores. These measures can be likened to certain situations and linked to particular areas of executive difficulty meaning that the full complement of scores for these tasks can yield rich information about an individual's relative strengths and difficulties. For example, the total time paused composite score for the School Bag Task might provide information on an individuals' cognitive control, while their planning efficiency score might yield additional information on their prospective memory and planning skills. This feature makes these tasks potentially more clinically useful as they can provide a profile of abilities allowing for targeted interventions.

2.5.3 Conclusion

The Eco-TED Consent Form Test, School Bag and Lego Tasks were developed and piloted. A floor effect may have obscured any group differences in Consent Form Test scores, but significant differences in the level of impairment were found for the School Bag and Lego Tasks, employing a provisional scoring system devised using the data collected. The validity of this scoring system will need to be further tested in a replication sample. It is difficult to draw conclusions on the specific executive domains involved but it seems likely that at least planning, prospective memory, working memory, attention, inhibition and interference control are implicated.

Verisimilitude was prioritised throughout the development of the tasks but unfortunately veridicality appears to be an area of weakness for the Eco-TED. Equally, test-retest reliabilities were low though to some extent this may relate to the real-life nature of the tasks. Further exploration of the Eco-TED's construct validity is also required.

2.6 References

- American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders fifth edition. *Arlington: American Psychiatric Publishing*, 280.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th edn). Englewood Cliffs, NJ: Prentice Hall.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, a, Swettenham, J., Wheelwright, S., & Drew, a. (2000). A screening instrument for autism at 18 months of age: a 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(6), 694–702.
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet*, 368(9531), 210–215.
- Baron-Cohen. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, 6(6), 248–254.
- Baron-Cohen, S., Wheelwright, S., Stone, V. E., & Rutherford, M. (1999). A mathematician, a physicist and a computer scientist with Asperger syndrome: Performance on folk psychology and folk physics tests. *Neurocase*, 5(6), 475–483.
- Baron, I. S. (2007). Review of Behavioural Assessment of the Dysexecutive Syndrome for Children (BADSC). *Child Neuropsychology*, 13(6), 539–542.
- 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, 194–209.
- Burgess, P. W., Alderman, N., Wilson, B. A., Evans, J. J., & Emslie, H. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADs)* (Vol. 5). Bury St. Edmunds, UK: Thames Valley Test Company.
- Burgess, P. W., & Shallice, T. (1997). The Hayling and Brixton tests. *Bury St Edmunds, UK: Thames Valley Test Company Limited.*
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *JAMA*, 285(24), 3093–9.
- Cohen, J. (1992). Quantitative Methods in Psychology. *Psychological Bulletin*, 112(1), 155–159.
- Duncan, J. (1986). Disorganisation of behaviour after frontal lobe damage. *Cognitive Neuropsychology*, 3(3), 271–290.
- Emslie, H., Wilson, F., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2003). *Behavioural Assessment of the Dysexecutive Syndrome for Children (BADs-C)*. London: Harcourt Assessment/The Psychological Corporation.
- Engel-Yeger, B., Josman, N., & Rosenblum, S. (2009). Behavioural Assessment of the Dysexecutive Syndrome for Children (BADs-C): An examination of construct validity. *Neuropsychological Rehabilitation*, 19(5), 662–676.
- Frith, U. (1989). *Autism: Explaining the enigma*. (Vol. 7). Oxford: Blackwell.
- Frith, U. (2003). *Autism: Explaining the enigma*. (2nd editio). Oxford: Blackwell.
- Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences*, 13(2), 74–82.
- Gioia, G. a, Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, 6(3), 235–238.

- Goodman, A., & Goodman, R. (2009). Strengths and difficulties questionnaire as a dimensional measure of child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(4), 400–3.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(5), 581–586.
- Happé, F. G. E. (1994). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 35(8), 1461–1471.
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test Manual: Revised and expanded*. New York: Psychological Assessment Resources.
- Henry, L., & Bettenay, C. (2010). The Assessment of Executive Functioning in Children. *Child and Adolescent Mental Health*, 15(2), 110–119.
- Hill, E. L. (2004a). Evaluating the theory of executive dysfunction in autism. *Developmental Review*, 24(2), 189–233.
- Hill, E. L. (2004b). Executive Dysfunction In Autism. *Trends in Cognitive Sciences*, 8(1), 26–32.
- Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia*, 44(14), 2822–2835.
- Hill, E. L., & Russell, J. (2002). Action memory and self-monitoring in children with autism: self versus other. *Infant and Child Development*, 11(2), 159–170.
- Homack, S. R., & Reynolds, C. R. (2007). *Essentials of assessment with brief intelligence tests*. Psychological Assessment.

- Irby, S. M., & Floyd, R. G. (2013). Test Review: Wechsler Abbreviated Scale of Intelligence. *Canadian Journal of School Psychology, 28*(3), 295–299.
- Kenworthy, L., Yerys, B. E., Anthony, L. G., & Wallace, G. L. (2008). Understanding Executive Control in Autism Spectrum Disorders in the Lab and in the Real World. *Neuropsychology Review, 18*(4), 320–338.
- Lai, M. C., Lombardo, M. V., Pasco, G., Ruigrok, A. N. V., Wheelwright, S. J., Sadek, S. A., ... Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PLoS ONE, 6*(6).
- Leitner, Y. (2014). The Co-Occurrence of Autism and Attention Deficit Hyperactivity Disorder in Children – What Do We Know? *Frontiers in Human Neuroscience, 8*, 268.
- Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., ... Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849–861.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders, 19*(2), 185–212.
- Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14*(5), 749–61.
- Monchi, O., Petrides, M., Strafella, A. P., Worsley, K. J., & Doyon, J. (2006). Functional role of the basal ganglia in the planning and execution of actions.

- Annals of Neurology*, 59(2), 257–264.
- Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex*, 12(4), 313–324.
- Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, Concurrent and Construct Validity. *Neuropsychological Rehabilitation*, 10(1), 33–45.
- Nydén, A., Hagberg, B., Goussé, V., & Rastam, M. (2011). A cognitive endophenotype of autism in families with multiple incidence. *Research in Autism Spectrum Disorders*, 5(1), 191–200.
- Odhuba, R. a, van den Broek, M. D., & Johns, L. C. (2005). Ecological validity of measures of executive functioning. *The British Journal of Clinical Psychology / the British Psychological Society*, 44(Pt 2), 269–278.
- Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R. B., Martin, A., & Wallace, G. L. (2012). WISC-IV profile in high-functioning autism spectrum disorders: Impaired processing speed is associated with increased autism communication symptoms and decreased adaptive communication abilities. *Journal of Autism and Developmental Disorders*, 42(5), 655–664.
- Ozonoff, S. (1995). Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism. *Neuropsychology*, 9(4), 491–500.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 32(7), 1081–1105.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 37(I), 51–87.

- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ (Clinical Research Ed.)*, 316(7139), 1236–1238.
- Pullinger, J. (2016). *Developing and Piloting a New Measure of Executive Functioning for Children with Autism Spectrum Disorder (ASD)*. University College London.
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, 27(2), 224–260.
- Russell, J., & Hill, E. L. (2001). Action-monitoring and intention reporting in children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 42(3), 317–328.
- Russo, N., Flanagan, T., Iarocci, G., Berringer, D., Zelazo, P. D., & Burack, J. A. (2007). Deconstructing executive deficits among persons with autism: Implications for cognitive neuroscience. *Brain and Cognition*, 65(1), 77–86.
- Sergeant, J. a., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? *Behavioural Brain Research*, 130(1–2), 3–28.
- Sinzig, J., Morsch, D., Bruning, N., Schmidt, M. H., & Lehmkuhl, G. (2008). Inhibition, flexibility, working memory and planning in autism spectrum disorders with and without comorbid ADHD-symptoms. *Child and Adolescent Psychiatry and Mental Health*, 2(1), 4.
- Skuse, D. H., James, R. S., Bishop, D. V, Coppin, B., Dalton, P., Aamodt-Leeper, G., ... Jacobs, P. a. (1997). Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function. *Nature*, 387(6634), 705–8.
- Skuse, D. H., Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits : heritability , reliability and validity of the Social and Communication Disorders

- Checklist service Measuring autistic traits : heritability , reliability and validity of the Social and Communication Disorders Checklist, 568–572.
- Skuse, D. H., Mandy, W., Steer, C., Miller, L. L., Goodman, R., Lawrence, K., ... Golding, J. (2009). Social communication competence and functional adaptation in a general population of children: preliminary evidence for sex-by-verbal IQ differential risk. *Journal of the American Academy of Child and Adolescent Psychiatry*, 48(2), 128–137.
- Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., & Place, M. (2004). The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43(5), 548–558.
- Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychological Research*, 63(3–4), 289–298.
- Stuss, D. T., Alexander, M. P., Floden, D., Binns, M. A., Levine, B., McIntosh, A. R., ... Hevenor, S. J. (2009). Fractionation and Localization of Distinct Frontal Lobe Processes: Evidence from Focal Lesions in Humans. In *Principles of Frontal Lobe Function*.
- Towgood, K., Meuwese, J., Gilbert, S., Turner, M., & Burgess, P. (2009). Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. *Neuropsychologia*, 47, 2981–2988.
- Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*.
- Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence - Second Edition*. The Psychological Corporation: Harcourt Brace & Company. New York, NY.
- Weiner, I. B. (2003). Assessment Psychology. In *Handbook of Psychology: Volume*

1 (pp. 279–302).

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on “open-ended” executive function tests in Autism. *Autism Research*, 2(3), 138–147.

Willcutt, E. G., Sonuga-Barke, E. J. S., Nigg, J. T., & Sergeant, J. A. (2008). Recent developments in neuropsychological models of childhood psychiatric disorders. *Advances in Biological Psychiatry*, 24, 195–226.

Wood, R. L., & Liossi, C. (2006). The ecological validity of executive tests in a severely brain injured sample. *Archives of Clinical Neuropsychology*, 21(5), 429–437.

Part 3: Critical Appraisal

3.1 Introduction

This appraisal sets out insights I have gained through the process of carrying out the current research study and literature review. First, general reflections on conducting research with individuals with Autism Spectrum Disorder (ASD) are discussed. My observations on the process of neuropsychological test development are then considered, followed by comments on the experience of supervising research assistants. This appraisal closes with a discussion on the future of the Ecologically-Valid Test of Executive Dysfunction (Eco-TED) and a brief conclusion on the appraisal as a whole.

3.2 Carrying out research with individuals with ASD

Perhaps unusually for a doctoral student, I opted to carry out research in an area I had limited knowledge of. Although I gained some experience of working with children with ASD as a part-time research assistant prior to clinical training, my understanding of ASD was almost entirely theoretical before I undertook these projects. Doing so provided me with ample opportunity to gain a better understanding of the nature of this condition and the implications this has for how research into this area is carried out. Some key ideas are discussed below.

3.2.1 Confounding variables

The process of carrying out research into ASD has impressed upon me the complexity of this condition and the great number of participant-related variables that can affect performance on tasks used in research, particularly those assessing executive functioning. In the literature review, we specifically investigated whether age affected planning performance and although we did not find evidence for age-related effects, other researchers have; both with regard to planning (Happé, Booth,

Charlton, & Hughes, 2006) and other executive functions (Huizenga, Ingmar, & Conor, 2011). When it came to planning and interpreting the findings of the meta-analysis and research study, I was struck by the number of potential confounds, aside from age, which had to be considered. For example, it was important to consider the possible impact of comorbid attention deficit hyperactivity disorder (ADHD). With regard to the review, it was not always possible to get clarification on the clinical characteristics of the ASD participants who took part in the studies and therefore it was necessary to consider whether comorbid ADHD might have somehow influenced our findings. With regard to our research study, we decided to include participants with comorbid ADHD if they wished to take part. This was partly because, in the context of the high incidence of comorbid ADHD in ASD (Leyfer et al., 2006), we were concerned about the impact limiting our sample in this way would have on our ability to recruit participants. In addition, it was also thought that including children with comorbid ADHD would give our findings greater generalisability. This decision was therefore made for valid reasons, but high numbers of children in our ASD group who also had ADHD diagnoses would of course have jeopardised our ability to draw conclusions on ASD specifically.

IQ in ASD presents a similar dilemma. Almost all of the studies included in the review focused on children with an IQ within the normal ranges. In a population where at least 25% have an intellectual disability (Chakrabarti & Fombonne, 2001), this had obvious implications for the generalisability of their, and therefore our, findings. Nonetheless, when it came to deciding upon our own inclusion and exclusion criteria for the research study, we too decided to limit our sample to individuals with an IQ > 70. In this way our sample was not representative of the ASD population, but it was decided that IQ represented a potential confound and that

efforts should therefore be made to limit its effects as much as possible.

The experience of carrying out research into ASD has therefore shown me that a number of participant-related factors are involved in performance on research tasks aiming to measure aspects of executive functioning. More importantly, it has shown me that this may result in researchers restricting samples in ways that undermine the relevance of the research. Alternatively, it might mean including a wider range of individuals at the risk of producing findings that can only be interpreted very conservatively. Trade-offs may therefore often have to be made in order to carry out research in this area.

3.2.2 Heterogeneity among individuals with ASD

Individuals with ASD vary greatly in terms of their abilities and disabilities. Towgood, Meuwese, Gilbert, Turner, & Burgess (2009) propose that ASD is in fact characterised by this variability, both within and between individuals with ASD. This raises the issue of “averaging artefacts” (Shallice & Evans, 1978) where patterns detected at group level may not accurately describe any single member of the group. Therefore, even where significant group differences are found, it is likely that the observed effects at group level are not universal to the sample. Our exploration of individual variability in performance in the current research provided some support for this idea as our ASD group’s performance varied significantly more than our controls’. Although group comparisons are undoubtedly useful in advancing our understanding of the broad ASD phenotype, my experience of carrying out this research has made me aware of the value in taking a more individually oriented approach to researching ASD as a complement to more classic group research design.

3.2.3 Methodological challenges

The characteristics of ASD, comprising social communication, behavioural and sensory difficulties, mean that study methodology is particularly important, specifically with regard to choosing measures. In light of evidence suggesting that individuals perform less well on socially administered tasks than computerised tasks (Ozonoff, 1995), the literature review explored whether the mode of administration affected planning performance. Our findings provide tentative support to the idea that socially administered tasks present more of a challenge for individuals with ASD. The key aim of the research study was to develop and pilot a measure of executive functioning designed specifically for children with ASD. We prioritised ecological validity, taking a symptom-led, ‘bottom-up’ approach to task development and ensuring as much as possible that the tasks resembled the situations in which children with ASD might experience executive difficulties. In doing so we created a set of seven researcher-administered tasks. In this sense the Eco-TED may have placed additional socio-communication demands on our ASD participants, making it more difficult for them to perform well for reasons other than their executive difficulties. This may present a broader dilemma with regard to measure choice for research involving participants with ASD. Ecological validity is important as it has direct implications for the clinical usefulness of a measure (Burgess et al., 2006), but it seems that there is a conflict here between selecting or devising measures that resemble real-life, and those that reduce the number of demands placed on the individual, thus providing a potentially ‘purer’ measure of the construct in question. For example, using a computerised version of the School Bag Task (see Appendix V for task script) would of course have allowed us to discount the role of the social communication demands of having to complete the task with the help of a researcher,

but it would also clearly have reduced the task's resemblance to real-life situations.

Carrying out this research has therefore drawn my attention to trade-offs that may be inherent to task selection or design, specifically with regard to achieving good construct as well as ecological validity. Furthermore, it has led me to question whether these compromises are in fact necessary or useful. As mentioned above, ecologically valid tests should resemble everyday scenarios as much as possible. In real-life situations, it is unlikely that individuals with ASD are presented with challenges that place demands on them solely in terms of one set of difficulties. For example, it seems likely that executive-type problems would often be experienced alongside social communication difficulties. I therefore wonder whether striving to examine these types of difficulties in isolation for the sake of construct validity is an entirely worthwhile endeavour as it seems that this would then affect the clinical relevance of the measure in question. This of course has implications for the advancement of our understanding of specific types of executive functioning constructs through research, but in the context of the social communication, behavioural and sensory difficulties individuals with ASD experience, it may be that separating these problem areas is somewhat artificial.

3.2.4 A deficit narrative

A narrative that emphasises the cognitive disabilities and impairments of individuals with ASD pervades the literature on this condition. Indeed, even evidenced areas of strength among individuals with ASD have often been understood “as low-level by-products of high-level deficits”, rather than “direct manifestations” of cognitive ability (Dawson, Soulieres, Gernsbacher, & Mottron, 2007, p. 657). This is arguably a damaging narrative and is likely to have implications for public awareness of, and attitudes towards, ASD.

The approach taken in the empirical paper assessed deficits in our ASD participants' performance on the Eco-TED relative to typically developing controls'. By relying on a 'deficit-focussed' approach to cognitive assessment, this research, and other studies like it, may unhelpfully, though unintentionally, contribute to an unbalanced narrative that focuses solely on the cognitive difficulties associated with ASD. Such an approach may effectively foreclose the possibility of discoveries that support alternative narratives, as it is likely that researchers who set out to identify deficits will overlook evidence of cognitive strengths in their data. Upon reflection, this issue is particularly pertinent to the current research study where we focused on subnormal performance among ASD participants without thoroughly exploring supra-normal performance. An interest in impairments is, to a certain extent, clinically justifiable in the design of cognitive measures as identified areas of difficulty can provide a target for intervention. Nonetheless, it would be valuable to explore both ends of the spectrum of abilities, as well as the varying degrees of task success and failure, in future research into the properties of the Eco-TED. This would allow us to carry out research that is less biased by a deficit narrative in relation to ASD, and therefore produce findings that more accurately capture the abilities of the samples in question. This in turn would help to contribute to a necessary and important shift towards a more rounded and representative narrative at both a research and societal level.

3.3 Neuropsychological test development

Prior to carrying out the current research study, although I had gained experience of administering neuropsychological tests with a variety of clients, I had not been involved in the development of such a measure. I felt that this would be a

valuable and unique experience to gain whilst on training. Some of my reflections on this process are set out below.

3.3.1 Practicalities

Developing the tests that would go on to form part of the Eco-TED was perhaps the most challenging part of the research process, particularly in the earlier stages of their development. Although the trainee I conducted the research project with had more clinical experience of working with children with ASD, we were both inexperienced in neuropsychological task development and often found it challenging to translate our ideas into realistic plans. We learnt a great deal from our supervisors through this process, both of whom have experience of developing neuropsychological measures. It was important for us to hold in mind the long-term clinical applicability of the Eco-TED tasks: this would depend not only on their ecological validity and ability to detect executive difficulties in children with ASD, but also on more pragmatic issues such as cost and practicality. For example, with specific reference to the School Bag Task, our initial idea had been that children would select items for the school day from a large selection of pictures of objects on a table. Piloting indicated that this would not be feasible for either research or clinical purposes as a very large table surface was needed to spread all the pictures out and make them visible to the participant. Through discussion with our supervisors we were able to identify a solution which went on to be used in the study: an A2 sized poster which could be easily laid out on almost any surface depicting all the items required alongside distractor items. Small but important details of this kind frequently came up during the early stages of task development, and this lengthy and iterative part of the research process was characterised by collaboration with our supervisors.

3.3.2 Piloting and service user involvement

Upon reflection, there are some ways in which my limited experience of working with this client group influenced my approach to the research. For example, although I gave general consideration to how the children in both groups might experience the testing session, particularly as it would last about 90 minutes, I don't think I had a clear sense of how our ASD participants in particular would find the process. During the task development phase, our main priority was designing fun tasks that would capture the executive difficulties that children with ASD experience and I think in some ways this overshadowed my consideration of their social communication difficulties and how this might impact on our participants' experience. In hindsight, I think it would have been better to give these details more careful thought. Unfortunately we were only able to informally pre-pilot the Eco-TED with typically developing children because of restrictions in terms of our research contracts at Great Ormond Street Hospital, but pre-piloting the Eco-TED with children with ASD would have been very helpful. It might also have been helpful to consult parents of children with ASD. Although we received positive feedback from parents once their children had taken part in the study, taking time to gain a better sense of how our ASD participants would find the session would perhaps have permitted us to approach these sessions even more sensitively.

3.3.3 Taking a 'bottom-up' approach to task development

As mentioned above, we took a symptom-led approach to Eco-TED task development rather than working 'top-down' from theoretical executive functioning constructs. Principally, we hoped that this would support our aim with regard to ecological validity. Having now completed this research project, I feel I have a greater appreciation of the advantages of approaching task development this way.

When it came to interpreting the results of the Eco-TED pilot, it was interesting to consider which executive domains our findings might pertain to. For example, when we discovered that our ASD participants paused for less time than controls between hearing the instructions and starting to select the items they needed for the School Bag Task, I immediately wondered what this might be telling us in terms of executive functioning. It seemed to me that this behaviour related to impaired behavioural inhibition, but this was an idea borne purely out of the face validity of this measure. Holding in mind the risks of making assumed links between observed behaviours and particular executive functions (Geurts, Corbett, & Solomon, 2009), I was cautious about making definitive statements with regard to the executive domains involved. At the same time, I felt it was useful to hypothesise about which executive constructs were implicated as, although our findings are tentative, it seemed that relating them to broader questions about executive dysfunction in ASD would put them in context and therefore make them more valuable.

Taking a symptom-led approach to task development is worthwhile because it removes the need for making assumptions about the constructs involved and arguably provides more readily interpretable information on behaviour. However, not linking the behaviours we observe, albeit cautiously, to executive constructs may limit the contribution the research (or measure) can make to our understanding of executive dysfunction in ASD. This seems to present a quandary in neuropsychological test development as researchers may find themselves torn, as I did, between wanting to position their research in the wider context, and not wanting to make invalid interpretative leaps about their findings.

3.3.4 Developing instruments for use in specific populations

The Eco-TED was designed to capture the specific difficulties with planning and organisation that children with ASD experience. To this end, task development was led by data collected at the Social Communication Disorders Clinic using the 3Di (Skuse et al., 2004). Developing a measure designed specifically for use with this client group was one of the central aims of the research study. The rationale was that this would increase the validity of the Eco-TED as evidence suggests that tasks not designed specifically for individuals with ASD may yield less valid measurements of their cognitive abilities. For example, Dawson et al. (2007) found that their ASD participants' scores were on average 30 percentile points higher when assessed using the Raven's Progressive Matrices (Raven, Raven, & Court, 1998) than when assessed using Wechsler scales of intelligence. Developing the Eco-TED specifically for children with ASD therefore seemed particularly important.

However, if we consider those Eco-TED items and sub-measures which show relatively more promise in terms of being able to differentiate between our ASD and control groups (e.g. item 6 of the School Bag Task; Appendix V), we might question whether they would be suitable for use with other client groups. Although our aim was to focus on executive problems experienced by children with ASD specifically, it is likely that children with different difficulties and diagnoses would also be able to complete these parts of the Eco-TED. If future validation studies involving different groups of interest indicated that this would be appropriate, this would clearly increase and diversify the Eco-TED's applicability, making it a more useful measure.

3.3.5 Control Group Choice

The current meta-analysis and research study included control groups comprised of typically developing children. To a certain extent the decision to use

typically developing rather than clinical control groups was a pragmatic one. It was clear from preliminary literature searches that studies using typically developing control groups were more prevalent and that focusing on these would therefore improve the quality of the meta-analysis in terms of the pooled volume of the data. Similarly, recruiting a suitable clinical control group for the purposes of the empirical paper would have been difficult within the constraints of doctoral research, particularly in terms of gaining access to suitable participants.

As well as the practical motivations outlined above, there was also a theoretical rationale for selecting a typically developing comparison group in both studies. Doing so allowed us to characterise the executive difficulties children with ASD have compared to a 'baseline' set by typically developing children. Although this approach has its limitations, this is a helpful comparison to make as children with ASD are very often required to function in settings oriented towards typically developing children. Such an approach therefore allows us to develop an understanding of how children with ASD might function in real-world scenarios, compared to neurotypical children of a similar age.

Selecting an appropriate clinical control group would have offered a different set of advantages. Although typically developing controls provide a useful benchmark against which to map performance, they do not allow researchers to control for other factors that may play an important role. For example, as neurotypical children obviously have no experience of having had a developmental disorder since a young age, or having had both their schooling and peer relations disrupted, it is impossible for us to rule out the impact of these factors on the performance of our ASD participants in both the meta-analysis and research study (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). This therefore makes it

impossible to say with certainty whether any detected executive difficulties are primary to ASD, or rather a product of something else, such as the presence of a childhood-onset neurodevelopmental disorder. Choosing a clinical control group comprised of children with a diagnosis of, for example, Tourette Syndrome, which shares many features with ASD, would have allowed us to control for some of these important factors (Baron-Cohen et al., 1997). This would undoubtedly be a useful exercise in future meta-analyses or research involving the Eco-TED.

3.4 Supervision of research assistants

Managing a team of three research assistants who contributed to data collection for the research project was my first experience of acting as a supervisor. This may have had an impact on the research project and there are a number of things I would do differently in future. Before discussing these, I will consider the ways in which I feel I was successful.

3.4.1 Successes

Most importantly, perhaps because of my pre-training research experience, I was able to arrange and manage the large participant timetable on behalf of the team. The research assistants were an invaluable resource, but I think that my own role in organising their and my time effectively was important in the success of data collection. We worked well as a team and were ultimately able to recruit and test the number of participants we set out to.

Through the research post I held prior to training, I also gained experience of completing a long battery of tasks in a research context. Among other things, I learnt to balance the needs of the participant with the research agenda, and gained a good sense of how to pace sessions in a way that keeps participants engaged and

motivated. These were valuable skills and I was able to discuss and demonstrate these ideas with the research assistants, both through observation and ad hoc supervision. I based my approach to training them on my own experience of being trained as a research assistant and this was to some extent quite successful.

3.4.2 Areas for improvement

The addition of research assistants to our team was a relatively late development in the research process meaning that there was a limit to how much time could be spent training and observing them. Though, as mentioned above, I made efforts to train the team thoroughly, it would have been better to be able to commit more time to this. Having conducted two observations for each research assistant, I had confidence in their abilities, but it might also have been helpful to continue to observe them over time, for example by using audio or video recordings of the sessions. This would have allowed me to ensure that the standard of administration was being maintained and that it was also consistent between researchers. Finding a way to assess inter-rater reliability would have allowed me to assess this empirically. In light of the various demands of the project at the time, I judged this not to be necessary because of the very positive impression I gained of the standard of administration. However the approach suggested above would have undoubtedly been more rigorous.

It is possible that the relatively short period of training the research assistants went through had an impact on our findings. Through administering the Eco-TED myself and observing others in testing sessions, my impression is that it was relatively easy to use. A great deal of detail went into the task script, making it fairly self-explanatory, but it remained quite readable and clear. However, tasks such as the Six Part Test from the Behavioural Assessment of the Dysexecutive Syndrome in

Children (BADs-C; Emslie, Wilson, Burden, Nimmo-Smith, & Wilson, 2003) require more skill and, from experience, a great deal more practice. Interestingly, although the control group for this study was tested mainly by the current author, the ASD group was predominantly tested by research assistants. Our BADs-C test results were not as expected for the ASD group and it therefore seems possible that difficulties in administration were an issue.

3.5 The future of the Eco-TED

The development and pilot of the Eco-TED was successful in that we found that when our empirically derived deficit scoring method was applied, the School Bag and Lego Tasks were able to differentiate between our ASD and control groups. It therefore seems that it would be worth taking the development of these two tasks further. My recommendations are set out below.

3.5.1 A larger sample

The findings of our pilot need to be replicated using a larger sample before we can draw more definitive conclusions on the properties of the Eco-TED. Focusing on the School Bag and Lego Tasks, using a replication sample would provide an opportunity to improve certain aspects of these tests. For example, it would be helpful to improve the test-retest reliabilities of these tests, as well as the degree to which they correlate with everyday outcome variables. Using a larger sample would also permit us to further test the validity of our empirically derived scoring method and additionally would allow us develop a set of norms for this purpose based on a larger group of controls. Finally, with increased statistical power we could also carry out a factor analysis as a way of seeing which of the Eco-TED outcome variables group together. This would act as an ideal complement to our

‘bottom-up’ approach to task development as it would help us to make meaningful links between the behaviours observed and particular executive functions.

3.5.2 Item 6 of the School Bag Task

When we looked at the individual items of the School Bag Task, item 6 was the only one that could distinguish between controls and ASD participants. It seems that this item is therefore more potent than the other six School Bag Task items (see section 2.5.1.2 for a discussion). It would therefore be useful to introduce more items resembling item 6 into the School Bag Task in any further research as this might well increase the discriminative power of this task.

3.5.3 Questionnaire development

Klinger & Renner (2000) have emphasised the need for an approach that combines performance measures with parental reports in ASD research. It would therefore be helpful to devise a parent-report questionnaire that could ultimately be used in conjunction with the Eco-TED. This questionnaire would be comprised of items that correspond to the behaviours assessed by the Eco-TED tasks. Dywan & Segalowitz (1996) found that pairing the Behavioural Assessment of the Dysexecutive Syndrome (BADs; Burgess, Alderman, Wilson, Evans, & Emslie, 1996) and the Dysexecutive Questionnaire (DEX; Burgess et al., 1996) improved the validity of the former. The DEX is part of the extended BADs battery and is a questionnaire-based informant-report measure. We might therefore hope to accrue a similar advantage by pairing the Eco-TED with a parent-report questionnaire.

3.6 Conclusion

Through the process of carrying out this research, I have been struck by the complexity of ASD, and the dilemmas this presents in terms of study design.

Through collaboration with my supervisors, I have also learnt a great deal about the practicalities of neuropsychological test development and the value of taking a symptom-led approach to this, although this is not without its limitations. Supervising a small team of research assistants for the first time was a valuable experience and would help to guide me in future research. Overall, conducting this research project has given me a greater awareness of the compromises that often have to be made for the purposes of research.

The Eco-TED School Bag and Lego Tasks merit further investigation. This could involve replication with a larger sample, the development of new test items for the School Bag Task, and further validation of our empirically derived scoring method. Finally, the development of a parent-report questionnaire for use alongside the Eco-TED could help to improve its validity.

3.7 References

- Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another Advanced Test of Theory of Mind: Evidence from Very High Functioning Adults with Autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(7), 813–822.
- 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, 194–209.
- Burgess, P. W., Alderman, N., Wilson, B. A., Evans, J. J., & Emslie, H. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADS)* (Vol. 5). Bury St. Edmunds, UK: Thames Valley Test Company.
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *JAMA*, 285(24), 3093–9.
- Dawson, M., Soulieres, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science*, 18(8), 657–662.
- Dywan, J., & Segalowitz, S. J. (1996). Self- and family ratings of adaptive behavior after traumatic brain injury: Psychometric scores and frontally generated ERPs. *The Journal of Head Trauma Rehabilitation*, 11(2), 79–95.
- Emslie, H., Wilson, F., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2003). *Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C)*. London: Harcourt Assessment/The Psychological Corporation.
- Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences*, 13(2), 74–82.

- Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61*(1), 25–39.
- Huizenga, H. M., Ingmar, V., & Conor, D. (2011). Testing overall and moderator effects in random effects meta-regression. *British Journal of Mathematical & Statistical Psychology, 64*(1), 1–19.
- Klinger, L. G., & Renner, P. (2000). Performance-based measures in autism: implications for diagnosis, early detection, and identification of cognitive profiles. *J Clin Child Psychol, 29*(4), 479–492.
- Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., ... Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*(7), 849–861.
- Ozonoff, S. (1995). Reliability and Validity of the Wisconsin Card Sorting Test in Studies of Autism. *Neuropsychology, 9*(4), 491–500.
- Raven, J., Raven, J., & Court, J. (1998). *Raven manual: Section 3. Standard progressive matrices*. Oxford, England: Oxford Psychologists Press.
- Shallice, T., & Evans, M. E. (1978). The Involvement of the Frontal Lobes in Cognitive Estimation. *Cortex, 14*(2), 294–303.
- Skuse, D., Warrington, R., Bishop, D., Chowdhury, U., Lau, J., Mandy, W., & Place, M. (2004). The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(5), 548–558.
- Towgood, K., Meuwese, J., Gilbert, S., Turner, M., & Burgess, P. (2009).

Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. *Neuropsychologia*, 47, 2981–2988.

Appendix I

Critical Appraisal Skills Programme (CASP, 2014) checklist for case control studies



11 questions to help you make sense of case control study

How to use this appraisal tool

Three broad issues need to be considered when appraising a case control study:

- **Are the results of the trial valid?** (Section A)
- **What are the results?** (Section B)
- **Will the results help locally?** (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically.

The first two questions are screening questions and can be answered quickly. If the answer to both is “yes”, it is worth proceeding with the remaining questions.

There is some degree of overlap between the questions, you are asked to record a “yes”, “no” or “can’t tell” to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

These checklists were designed to be used as educational tools as part of a workshop setting

There will not be time in the small groups to answer them all in detail!

©CASP This work is licensed under the Creative Commons Attribution - NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/> www.casp-uk.net

(A) Are the results of the study valid?

Screening Questions

1. Did the study address a clearly focused issue? Yes Can't tell No

HINT: A question can be focused in terms of

- The population studied
- The risk factors studied
- Whether the study tried to detect a beneficial or harmful effect?

2. Did the authors use an appropriate method to answer their question? Yes Can't tell No

HINT: Consider

- Is a case control study an appropriate way of Answering the question under the circumstances? (Is the outcome rare or harmful)
- Did it address the study question?

Is it worth continuing?



Detailed questions

3. Were the cases recruited in an acceptable way?

Yes

Can't tell

No

HINT: We are looking for selection bias which might compromise validity of the findings

- Are the cases defined precisely?
- Were the cases representative of a defined population? (geographically and/or temporally?)
- Was there an established reliable system for selecting all the cases
- Are they incident or prevalent?
- Is there something special about the cases?
- Is the time frame of the study relevant to disease/exposure?
- Was there a sufficient number of cases selected?
- Was there a power calculation?

4. Were the controls selected in an acceptable way?

Yes

Can't tell

No

HINT: We are looking for selection bias which might compromise The generalisibility of the findings

- Were the controls representative of defined population (geographically and/or temporally)
- Was there something special about the controls?
- Was the non-response high? Could non-respondents be different in any way?
- Are they matched, population based or randomly selected?
- Was there a sufficient number of controls selected?

5. Was the exposure accurately measured to minimise bias?

Yes

Can't tell

No

HINT: We are looking for measurement, recall or classification bias

- Was the exposure clearly defined and accurately measured?
- Did the authors use subjective or objective measurements?
- Do the measures truly reflect what they are supposed to measure? (Have they been validated?)
- Were the measurement methods similar in the cases and controls?
- Did the study incorporate blinding where feasible?
- Is the temporal relation correct? (Does the exposure of interest precede the outcome?)

6. (a) What confounding factors have the authors accounted for?

List:

HINT: List the ones you think might be important, that the author missed.

- Genetic
- Environmental
- Socio-economic

(b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes

Can't tell

No

HINT: Look for

- Restriction in design, and techniques e.g. modelling stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

7. What are the results of this study?

HINT: Consider

- What are the bottom line results?
- Is the analysis appropriate to the design?
- How strong is the association between exposure and outcome (look at the odds ratio)?
- Are the results adjusted for confounding, and might confounding still explain the association?
- Has adjustment made a big difference to the OR?

(B) What are the results?

8. How precise are the results?

How precise is the estimate of risk?

HINT: Consider

- Size of the P-value
- Size of the confidence intervals
- Have the authors considered all the important variables?
- How was the effect of subjects refusing to participate evaluated?

9. Do you believe the results?

Yes

No

HINT: Consider

- Big effect is hard to ignore!
- Can it be due to chance, bias or confounding?
- Are the design and methods of this study sufficiently flawed to make the results unreliable?
- Consider Bradford Hills criteria (e.g. time sequence, dose-response gradient, strength, biological plausibility)

(C) Will the results help locally?

10. Can the results be applied to the local population?

Yes

Can't tell

No

HINT: Consider whether

- The subjects covered in the study could be sufficiently different from your population to cause concern
- Your local setting is likely to differ much from that of the study
- Can you quantify the local benefits and harms?

11. Do the results of this study fit with other available evidence?

Yes

Can't tell

No

HINT: Consider all the available evidence from RCT's, systematic reviews, cohort studies and case-control studies as well for consistency.

Remember

One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making.

However, for certain questions observational studies provide the only evidence.

Recommendations from observational studies are always stronger when supported by other evidence.

Appendix II

Ethical approval document

London - Westminster Research Ethics Committee

4 Minshull Street
Manchester
M1 3DZ

Telephone: [REDACTED]

21 September 2015

Dr William Mandy
University College London
Research Department of Clinical, Educational and Health Psychology
Gower Street
London
WC1E 6BT

Dear Dr Mandy

Study title: Validating a new ecologically valid measure of executive functioning for children with autism spectrum disorder (ASD)
REC reference: 15/LO/1332
Protocol number: N/A
IRAS project ID: 170531

Thank you for your submission of 11 September 2015, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair and Dr Yash Patel.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager, Ms Rachel Katzenellenbogen, [REDACTED]

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact Catherine Blewett [REDACTED], the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Letters of invitation to participant [Letter to parents (clinical) V3]	V3	10 April 2015
Letters of invitation to participant [Letter to parents (control) V4]	V4	10 April 2015
Non-validated questionnaire [Task Script V2]	V2	30 June 2015
Other [Summary CV for student 2]		
Other [D.Skuse (PI) CV]		
Other [Data protection confirmation]	1	10 April 2015
Other [Updated Insurance Certificate 2015-16]	2	09 September 2015
Other [Evidence of Peer Review]	1	09 December 2014
Other [Invitation letter to headteachers]	1	10 September 2015
Other [Response Point 2. Provisional Opinion]	1	11 September 2015
Participant consent form [Consent Clinical]	6	10 May 2015
Participant consent form [Consent (control)]	5	10 May 2015
Participant information sheet (PIS) [Child Information Sheet V3]	3	08 September 2015
Participant information sheet (PIS) [Parent Info Sheet Consent Clinical Test of EF V7]	7	08 September 2015
Participant information sheet (PIS) [Parent Info Sheet Consent Control Test of EF V6]	6	08 September 2015
REC Application Form [REC_Form_15072015]		15 July 2015
Research protocol or project proposal [Research Protocol V3]	V3	10 April 2015
Summary CV for Chief Investigator (CI) [Summary CV CI]		
Summary CV for student [Eleonore Bristow CV]		
Summary CV for supervisor (student research) [PBurgess Summary CV]		
Validated questionnaire [BRIEF]	1	30 June 2015
Validated questionnaire [BADS Tasks]	1	30 June 2015
Validated questionnaire [SCDC]	1	30 June 2015
Validated questionnaire [SDQ]	1	30 June 2015
Validated questionnaire [VABS-II]	1	30 June 2015
Validated questionnaire [Description WASI-II]	1	30 June 2015

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance

on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:

<http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/>

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at

<http://www.hra.nhs.uk/hra-training/>

15/LO/1332	Please quote this number on all correspondence
-------------------	---

With the Committee's best wishes for the success of this project.

Yours sincerely



Dr Alan Ruben
Chair



Enclosures: "After ethical review – guidance for researchers"

Copy to: Ms Smaragda Agathou, Joint Research Office
Mr Elliott Dickens, Great Ormond Street Hospital

Appendix III

Information sheets and consent forms

Information sheet and consent form 1 – for parents/carers of typically developing children

Information sheet and consent form 2 – for parents/carers of children with Autism Spectrum Disorder (ASD)

Information sheet and assent form – for both typically developing children and children with ASD

Information sheet and consent form for parents/carers of typically developing

Children (1)

PARENT/GUARDIAN INFORMATION SHEET AND CONSENT FORM OF TYPICALLY DEVELOPING ADOLESCENTS AGED 8-12 YRS

Developing a Measure of Planning and Organisation for Children with ASD

We would like to invite you and your child to take part in our research study. Before you and your child decide whether you would like to take part, it is important for you to know why the research is being done and what it will involve. Please take time to read this information sheet carefully and discuss it with others if you wish. If there is anything that is not clear, or if you would like more information, please do not hesitate to contact us.

Why is the study being done?

Please note that we are contacting you because we are keen to recruit comparison children who do not have a diagnosis of Autistic Spectrum Disorder (ASD). These children will form part of our control group.

Executive functioning is a term used to describe the many tasks our brain performs that are necessary to think, act, and solve problems. It includes tasks that help us learn new information, remember and retrieve information we've learned in the past, and use this information to solve problems of everyday life.

Children who find these things difficult can struggle in different aspects of their life. For the purpose of our study we are particularly interested in looking at thinking, learning and planning in children with ASD. There are currently a number of tests that aim to assess these skills but the problem with those already available is that they have not been specifically designed for children with ASD. For this reason they can miss some of the everyday difficulties that are seen in individuals with the diagnosis. Through our research we are hoping to develop a new test that more accurately assesses thinking, learning and planning in children with ASD.

We need a control group so that we can compare how well the children in the control group do on the test compared to the children with ASD. If the children in the control group do better on the test then we will know that our test is good at differentiating between children with and without the disorder. That is why we'd like your child to take part.

What will happen if we take part?

If you agree to take part in this research, your child will be seen by either [REDACTED] or [REDACTED] (study researchers) at school or at home, depending on which location is more convenient for you. The session will last for a maximum of two hours. During that time, your child will do some games and puzzles that look at how they think and process information. The games have been designed to be fun.

You will be asked to fill in some questionnaires about your child's behaviour, communication, strengths and weaknesses. These are simple parent-report questionnaires which are widely used and should take you no longer than one hour to complete.

A small number of children will be asked to take part in the games and puzzles for a second time. This shorter follow-up session will take place around a month after the first visit and will take no longer than one hour. Only your child will need to take part in this session.

An information sheet for your child has been provided. Please talk about the study with your child. We will also make sure that your child understands what he/she will be doing and give them an opportunity to ask any questions that they may have.

As a small thank you for taking part in our study we will offer your child a £5 voucher and enter them into a draw to win another £50 voucher.

What are the potential benefits?

We hope that our findings will help to develop a more reliable measure of executive functioning for children with ASD. There is likely to be no immediate benefit for the children taking part in the study, but we hope that their help will be beneficial to other children in the future.

Does my child have to take part in this study?

It is up to you and your child whether or not to take part in this study. We kindly ask you to complete the attached form and return it to your child's teacher indicating whether you would/would not like your child to take part. If you do decide to take part but later change your mind you are free to withdraw at any time without giving a reason. Even if you are happy for your child to take part, he or she will still decide for himself. It will be explained to your child that he/she can choose to withdraw from the study at any time, without giving a reason. We want to make sure that everyone is happy when taking part in our project.

Will taking part in this study remain confidential?

All information collected from you and your child during the course of this research will be kept strictly confidential. No one, other than the researchers involved in the study, will have access to your or your child's personal details or any of the information provided to the Service. This

information will be kept in locked cabinets and stored anonymously at University College London (UCL).

Who has reviewed this study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given a favourable opinion by the Westminster Research Ethics Committee.

What will happen to the results of the research?

The information collected from children with a diagnosis of ASD will be compared to a group of children without ASD to see whether the test of executive functioning is useful in differentiating between those with and without the diagnosis. The findings of the study will be written up by the researchers as part of two doctoral theses. However, names and other identifying information will be removed. The results of the study may be presented at national and international conferences and published in academic journals. Neither you nor your child will be personally identified in any reports or publications of the research. If you wish, a summary of the findings can be sent to you via post or email once the study is complete.

How to contact the researchers

If you have any further questions or would like assistance at any point during the study, please feel free to contact [REDACTED] or [REDACTED] at UCL on [REDACTED] [REDACTED] or email [REDACTED]. In the case of a complaint, please contact [REDACTED] via [REDACTED].

We are happy to talk through any questions with you.

Thank you for taking the time to read this information sheet.

Your help makes our research possible!

University College London holds insurance against claims from participants for harm caused by their participation in this clinical study. Participants may be able to claim compensation if they can prove that UCL has been negligent. However, if this clinical study is being carried out in a hospital, the hospital continues to have a duty of care to the participant of the clinical study. University College London does not accept liability for any breach in the hospital's duty of care, or any negligence on the part of hospital employees. This applies whether the hospital is an NHS Trust or otherwise.

Please tick (✓) appropriate box:

Yes, my child and I are happy to participate in this study (Version 5.0 10/05/15)

No, we do not want to participate in this study.

If Yes, please complete the following:

(Please initial box)

I have read the Information Sheet (Version 6, 08/09/2015).

I understand that I am free to withdraw my child from the study at any time without giving a reason.

I understand that my child is free to withdraw from the study at any time without giving on.

I give consent to be sent some questionnaires to complete regarding my child.

I am happy to be contacted for a second time to arrange a shorter follow-up session.

I have had the opportunity to ask any questions I wish to ask.

I have the contact details of the research team in case I have any queries in the future.

Child's Name: _____

Parent's Name: _____

Parent/Guardian Signature: _____

Date: _____

Researcher Signature: _____

Date: _____

Contact Details (these will remain confidential and only be used to send questionnaires and arrange a session to meet with your child):

Address: _____

Tel. No: _____

PLEASE PROVIDE THE FOLLOWING DETAILS IF YOU WOULD LIKE TO BE SENT A SUMMARY OF THE FINDINGS ONCE THE STUDY IS COMPLETED

Email: _____

Information sheet and consent form for parents/carers of children with ASD (2)

PARENT/GUARDIAN INFORMATION SHEET AND CONSENT FORM FOR CHILDREN WITH A DIAGNOSIS OF AUTISM

Developing a Measure of Planning and Organisation for Children with ASD

We would like to invite you and your child to take part in our research study. Before you and your child decide whether you would like to take part, it is important for you to know why the research is being done and what it will involve. Please take time to read this information sheet carefully and discuss it with others if you wish. If there is anything that is not clear, or if you would like more information, please do not hesitate to contact us.

Why is the study being done?

There has been lots of research that has suggested that people with Autism Spectrum Disorders (ASD) can have difficulties with executive functioning. This is a term used to describe the many tasks our brains perform that are necessary to think, act, and solve problems. Executive functioning includes tasks that help us learn new information, remember and retrieve information we've learned in the past, and use this information to solve problems of everyday life.

There are a number of tests currently available that aim to assess a child's executive functioning. The problem with these tests is that they have not been specifically designed for children with ASD. For this reason the tests can sometimes miss some of the everyday difficulties that are seen in individuals with the diagnosis. Through our research we are hoping to develop a new test that more accurately assesses these difficulties so that we can gain a better idea of how executive functioning is affected in those with ASD.

What will happen if we take part?

If you agree to take part in this research, your child will be seen by one of the study researchers at the Social Communications Disorder Clinic at Great Ormond Street Hospital, or at your home depending on what is most convenient for you. This meeting will last no longer than two hours. During that time, your child will do some games and puzzles that look at how they think and process information. The games have been designed to be fun.

You will be asked to fill in some questionnaires about your child's behaviour, communication and feelings. These are widely used and should take no longer than one hour to complete. These can be completed whilst your child takes part in the games and puzzles or in your own time.

In addition we will also ask for your permission to access some of the routine information collected as part of your child's assessment at the Social Communications Disorders clinic. This will include information on your child's diagnosis and their IQ score. If you give permission we will liaise directly with your child's care team to collect this information.

A small number of children will be asked to take part in the games and puzzles for a second time. This shorter follow-up session will take place around a month after the first visit and will take no longer than one hour. Only your child will need to take part in this session.

An information sheet for your child has been provided. Please talk about the study with your child. We will also make sure that your child understands what he/she will be doing and give them an opportunity to ask any questions that they may have.

As a small thank you for taking part in our study we will offer your child a £5 voucher and enter them into a draw to win another £50 voucher.

What are the possible disadvantages and risks of taking part?

Whilst we expect that most children will enjoy the puzzles and games, it is possible that some children may find them hard work or frustrating to complete. We will offer regular breaks and give your child the opportunity to stop at any time should this happen. They will also have the chance to talk to the researcher's about how they found taking part once finished.

What are the potential benefits?

We hope that our findings will help to develop a better measure of executive functioning for children with ASD. There is no immediate benefit for the children taking part in the study, but we hope that their help will be beneficial to other children in the future.

Does my child have to take part in this study?

It is up to you and your child whether or not you take part in this study. If you do decide to take part, you will be asked to sign a consent form. If you decide now, or at a later date, that you do not wish to participate in this research you are free to withdraw at any time without giving a reason. Even if you are happy for your child to take part, he or she will still decide for himself. It will be explained to your child that he/she can choose to withdraw from the study at any time, without giving a reason. We want to make sure that everyone is happy when taking part in our project. We would also like to stress that if you decide not to take part in the research; it will not in any way affect the care that your child receives.

Will taking part in this study remain confidential?

All information collected from you and your child during the course of this research will be kept strictly confidential. No one, other than the researchers involved in the study, will have access to your or your child's personal details or any of the information provided to the Service. This information will be kept in locked cabinets and stored anonymously at University College London (UCL).

Who has reviewed this study?

All research in the NHS is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given a favourable opinion by the Westminster Research Ethics Committee.

What will happen to the results of the research?

The information collected from children with a diagnosis of ASD will be compared to a group of children without ASD to see whether the test of executive functioning is useful in differentiating between those with and without the diagnosis. The findings of the study will be written up by the researchers as part of two doctoral theses. However, names and other identifying information will be removed. The results of the study may be presented at national and international conferences and published in academic journals. Neither you nor your child will be personally identified in any reports or publications of the research. If you wish, a summary of the findings can be sent to you via post or email once the study is complete.

How to contact the researchers

If you would like to take part please contact Rob Hickman (Recruitment Assistant) on 0207 405 9200 ext: 1432. Rob will then pass on your details to the researchers so that they may contact you to make arrangements. If you have any further questions or would like assistance at any point during the study, please feel free to contact [REDACTED] or [REDACTED] at UCL by email [REDACTED]. In the case of a complaint, please contact [REDACTED] via [REDACTED]. We are happy to talk through any questions with you.

Thank you for taking the time to read this information sheet.

Your help makes our research possible!

University College London holds insurance against claims from participants for harm caused by their participation in this clinical study. Participants may be able to claim compensation if they can prove that UCL has been negligent. However, if this clinical study is being carried out in a hospital, the hospital continues to have a duty of care to the participant of the clinical study. University College London does not accept liability for any breach in the hospital's duty of care, or any

negligence on the part of hospital employees. This applies whether the hospital is an NHS Trust or otherwise.

Please tick (✓) appropriate box:

Yes, my child and I are happy to participate in this study

No, we do not want to participate in this study.

If Yes, please complete the following:

(Please initial box)

I have read the Information Sheet (Version 7, 08/09/15).

I understand that I am free to withdraw my child from the study at any time without giving a reason.

I understand that my child is free to withdraw from the study at any time without giving on.

I give consent to be sent some questionnaires to complete regarding my child.

I am happy to be contacted for a second time to arrange a shorter follow-up session.

I have had the opportunity to ask any questions I wish to ask.

I have the contact details of the research team in case I have any queries in the future.

Child's Name: _____

Parent's Name: _____

Parent/Guardian Signature: _____

Date: _____

Researcher Signature: _____

Date: _____

Contact Details (these will remain confidential and only be used to send questionnaires and arrange a session to meet with your child):

Address: _____

Tel No: _____

PLEASE PROVIDE AN EMAIL ADDRESS IF YOU WOULD LIKE TO BE SENT A SUMMARY OF THE FINDINGS ONCE THE STUDY IS COMPLETED

Email: _____

Information sheet and assent form for both typically developing children and children with ASD

Can you help us?

Our names are [REDACTED] and [REDACTED] and we are looking for some young people to take part in our study. This page tells you a bit about our study and we would be really grateful if you could have a read and see if you'd like to take part. If you are not sure about any of the words or have any questions please ask us or talk about it with a member of your family.

What is the study about?

We know that everyone thinks differently. Some people find it hard to learn new things and other people find it easy. Some people are good at solving puzzles whilst other people are good at telling stories. Most people have some things they are quite good at and others that they are not so good at.

We have designed some new games and puzzles that can help us to look at the different way young people think and do things. We would like you to give our games and puzzles a go to see if they work!



What will I need to do?

If you like the idea of taking part in our study then [REDACTED] or [REDACTED] will visit you at your home, school or at Great Ormond Street Hospital. You will get to have a go at some games and puzzles which we hope you'll find fun.

We'll also ask the person who takes care of you to fill in some questions about you. Things you are good at and things that you like doing.

As a small thank you for taking part in our study we will offer you a £5 voucher and enter you into a draw to win another £50 voucher.

Why ask me?

We are asking you because we want to test out our puzzles and games on young people who are between 8 and 12 years old.

What will it be like to take part?

We hope our games will be fun but sometimes you might find them a bit tricky. Not everyone will be able to finish them all. If you get tired or need a rest then you can ask to stop.

Do I have to take part?

No - it is up to you and the person who looks after you. If you do want to take part, we will ask you and your parent/carer to tick and sign a form. If you change your mind that's OK, you just have to tell us and you can stop at any time. You do not have to take part in this study.

Will anyone know how I do?



Our study is confidential. This means that no one will know how well you did in the puzzles and games.

Questions?

If you have any questions or would like to talk more about taking part you can ask to speak to [redacted] or [redacted].

- I know that I don't have to take part if I don't want to
- If I change my mind I can just tell my parent, [redacted] or [redacted]
- It's OK to ask my parent/carer some questions about me
- I am happy to take part in the games and puzzles twice if needed to

Please put a circle around No or Yes to tell us if you want to take part



No



Yes

Signed.....

Please print your name.....

Please give this form to your parent / carer as soon as possible

Appendix IV

Initial contact letters for parents

Initial contact letter 1 – for parents/carers of typically developing children

Initial contact letter 2 – for parents/carers of children with ASD

Initial contact letter for parents/carers of typically developing children (1)

Invitation for your child to take part in a study

Dear Parent / Guardian

We are carrying out some research and are looking to recruit children that do not have a diagnosis of Autistic Spectrum Disorder (ASD). The aim of the research is to validate a newly developed measure that hopes to accurately assess planning and organisation in children with ASD. We would like to invite you and your child to be part of our comparison group of children that do not have the disorder. An explanation of these terms and more detail about the study can be found on the enclosed information sheet.

This study will involve no more than two hours of yours and your child's time. Your child will be asked to do a number of games and puzzles that aim to assess how they think, learn and remember things. These are designed to be as fun as possible. You will be asked to complete some questionnaires about your child's behaviour, strengths and weaknesses.

A small number of children and parents will be contacted a second time to arrange a shorter follow-up meeting. Only your child will take part in this second visit and will repeat the puzzles and games which should take no longer than one hour

As a thank you your child will receive a £5 voucher and will also be entered into a draw to win one of two £50 vouchers.

We would be really grateful if you could spare some time to read through the attached information sheet and speak with your child about whether or not they would be happy to take part in the research. If you and your child decide to take part in the research then please return the enclosed consent form to your child's teacher and a researcher will contact you.

Please do not hesitate to contact us should you have any questions or require further information. Our contact details can be found on the information sheet attached.

Thank you for your time and for considering taking part in our research.

Yours Sincerely,

██████████

Trainee Clinical Psychologist
(Researcher)

██████████

Trainee Clinical Psychologist
(Researcher)

████████████████████

Clinical Psychologist
(Chief Investigator)

Initial contact letter for parents/carers of children with ASD (2)

Invitation for your child to take part in a study

Dear Parent / Guardian

We would like to invite you and your child to take part in some research that we are conducting at the Social Communication Disorders Clinic at Great Ormond Street Hospital. We are approaching you as you have previously given your permission to be contacted for research purposes.

Many children with Autism Spectrum Disorders (ASD) have difficulties planning and organising their behaviour. Currently there are no tests specifically designed to measure these difficulties in children with ASD. We are developing such a test, and want to find out whether the test works and is able to measure these things accurately. A more detailed explanation of the study can be found on the enclosed information sheet.

The study will involve no more than two hours of yours and your child's time, during which they will be asked to do a number of games and puzzles that aim to assess how they think, learn and remember things. These are designed to be as fun as possible. You will be asked to complete some questionnaires about your child's behaviour. This can be organised for a time and place of your choosing to minimise inconvenience to you and your child. .

A small number of children and parents will be contacted a second time to arrange a shorter follow-up meeting. Only your child will take part in this second visit and will repeat the puzzles and games which should take no longer than one hour.

As a thank you your child will receive a £5 voucher and will also be entered into a draw to win one of two £50 vouchers.

We would be grateful if you could spare some time to read through the attached information sheet and speak with your child about whether or not they would be happy to take part in the research. If you and your child are happy to take part then please contact [REDACTED] (Recruitment Assistant) on [REDACTED]. [REDACTED] will then pass on your details to the researchers so that they may contact you to make arrangements.

Please do not hesitate to contact us should you have any questions or require further information.

Thank you for your time and for considering taking part in our research.

Yours Sincerely,

[REDACTED]
Trainee Clinical Psychologist
(Researcher)

[REDACTED]
Trainee Clinical Psychologist
(Researcher)

[REDACTED]
Clinical Psychologist
(Chief Investigator)

Appendix V

Ecologically-Valid Test of Executive Dysfunction (Eco-TED) Consent Form Test,
School Bag and Lego Tasks script

Consent Form Test Part 1

Items required:

- 1 x child information and assent sheet
- Pen/pencil for signature
- Folder or file

At beginning of the testing session the participant signs a copy of the child information and assent sheet. Once the participant has signed the assent form/sheet, it is placed in a folder and positioned within view, on a flat surface (e.g. a table or chair) to the right or left of the examiner, no more than 2m away from the table on which testing is taking place.

Before you leave today, I need to give you this piece of paper [hold it up in front of the child]. I will put it in here to keep it safe for now [make sure child attends to paper being put in file and file being placed on nearby surface]. At the end of our session, but before I/you leave, it is important that you ask me for it.

Just to make sure I explained that properly, can you tell me what I have asked you to do at the end of the session?

[If child does not convey understanding that they must ask for the information sheet before they leave, the examiner must repeat task instructions]

School Bag Task

Items required:

- School bag task poster
- Sticky labels
- Stop watch
- Lesson prompt cards

Participants are asked to plan for activities at school the next day. They will be presented with a poster (placed on the table directly in front of them) depicting 81 objects, including those they need for the given school activities, as well as distractor items which they do not need. Sticky labels are used to show which items they have selected. After each trial, the examiner **removes the labels**.

The examiner gives the following instructions:

Look at this poster – it has lots of pictures on it. Let’s see if you can find some things on the poster.

Scanning trial 1:

Can you show me a jam jar?

If they point to the jar say:

Well done. When you want to pick something on the poster, you need to stick one of these on it to show me you have chosen it. Try sticking one on the jam jar.

[If they cannot find the jam jar within 60 seconds the examiner points it out, and proceeds to the next scanning trial]

Scanning trial 2:

Hand the child the sticker dispenser and say:

Now find the sunglasses.

[If child does not put the sticker on, remind them that they need to do this whenever they pick an item]

[If they cannot find the sunglasses within 60 seconds the examiner points them out, and proceeds to the next scanning trial]

Scanning trial 3:

Now find the Clock

[If child does not put the sticker on, remind them that they need to do this whenever they pick an item]

[If the child fails all three scanning items, do not proceed with the task]

Auditory memory trial 1:

Now I am going to ask you to find three things, and put stickers on them all. Please find a teddy bear, a rubber duck and some tennis balls.

[If participant cannot remember what they need to find do not prompt them. Instead, say: “I can’t tell you again. Just try to remember what I said, and if you can’t remember just guess”]

[If participant finds the three items within 120 seconds, proceed to the main task. Otherwise, attempt auditory memory trial 2]

Auditory memory trial 2:

Please find a scarf, some sun cream and a blue pencil.

[If participant cannot remember what they need to find do not prompt them. Instead, say: “I can’t tell you again. Just try to remember what I said, and if you can’t remember just guess”]

[If participant finds the three items within 120 seconds, proceed to the main task. Otherwise, discontinue task]

Main task:

Now we will play a game using the poster. I will ask you to imagine that you have certain lessons and activities at school tomorrow. You will need to choose things from the poster to bring to school for these lessons and activities.

You can pack as much or as little as you like, but only pack exactly what you will need for each activity. Do not pack anything that you won’t need.

I’ll have a go at the game to show you how it is played.

Examiner takes practice prompt card and puts it in front of them. They then say:

The card shows me I have a French lesson tomorrow. So...I need to choose from the poster what I will need.

Examiner takes their time looking over whole poster and then says:

I’m going to put a sticker on this French dictionary because I’ll need it for the French lesson [put sticker on]. I mustn’t put a sticker on this sun lotion because I won’t need it for the French lesson.

Now it’s your turn – but before you start do you have any questions about this game?

Remember, you can pack as much or as little as you like for any activity, but don’t pack things you won’t need for that activity.

If at any stage the participant selects items they would not need for the lesson in question i.e. if they select an item not listed in the “correct items” section of the score sheet, you may prompt them ONCE ONLY over the course of the whole task:

Remember, you can pack as much or as little as you like, but don’t pack things you won’t need.

⇒ **Discontinue rule:** Discontinue task if child is unsuccessful on 3 consecutive trials i.e. if they do not select at least one of the correct items specified on the score sheet within 120 seconds.

Item 1

Show the participant Card 1 (i.e. Science). Say:

Okay, you're now going to choose what you'll need for your Science lesson. In this lesson you're going to be looking at plant leaves using a magnifying glass. What do you need?

[If child does not use stickers record their answers anyway, but prompt them to use stickers for subsequent trials. For subsequent trials prompt the participant to use stickers as needed. Also prompt the participant to tell you when they are finished if they do not do this of their own accord.]

For all items record the correct and incorrect items packed. Also record the time taken for the participant to select their items, starting timing as soon as the instructions are complete. You may repeat the instructions once if the participant asks but start recording as soon as the instructions are complete. Also note whether the participant paused between hearing the instructions and starting to select their items (see score sheet).

Item 2

Show the participant Card 2 (i.e. break time). Say:

It's break time. A friend in your class is bringing in a cake to celebrate their birthday. They asked you to bring in some balloons for them. What do you need?

Item 3

Show the participant Card 3 (i.e. Art). Say:

You have Art. Your teacher will ask you to draw a ladybird. You are only allowed to use colouring pencils for this. The ladybird should look like a real ladybird as much as possible. What do you need?

Item 4

Show the participant Card 4 (i.e. break time). Say:

It's break time. You will go to a shop near school to buy a bottle of water. It will probably rain during break time. What do you need?

Item 5

Show the participant Card 5 (i.e. lunch). Say:

At school tomorrow, you will have a packed lunch from home. You're not allowed chocolate or crisps, but you need to have a sandwich and two items containing fruit. What do you need?

Item 6

Show the participant Card 6 (i.e. P.E.). Say:

It's P.E. You can choose between tennis and football. You'll need shoes, socks, shorts and a top: white for tennis, red for football. School provides everything else. What do you need to bring to school?

[If the participant asks/tries to prepare for both football and tennis, ask them to choose and prepare for one option only]

Item 7

Show the participant Card 7 (i.e. end of school day). Say:

After P.E., it's the end of the school day. You could go to the school library – you've got a world atlas to return, and you want to get a book about the rainforest out. You've also got a five pence fine to pay off. Or you could go to homework club. You'll be saving your answers on a computer, but need to bring your maths textbook and calculator. You should bring a drink too. What do you need?

[If the participant asks/tries to prepare for both the library and homework club, ask them to choose and prepare for one option only]

Lego Task

Items required:

- Two bowls of Lego
- Spare Lego pieces for demonstration/learning
- Cup
- 6 small animal figures (3 in each Lego bowl)
- Stop watch

In this task participants will be asked to sort and select Lego. They will first be given the instructions below and shown 2 bowls containing 65 pieces of Lego each and three small animals. *NB the bowls should contain a greater quantity of each piece than the amount requested so that participants need to recall how many pieces were asked for i.e. if asked to find one piece; the bowl should contain at least two of these pieces.* Say:

In a moment I'd like you to help me find some Lego pieces that are hidden in these bowls.

I want you to find the following pieces as quickly as you can:

- ***A piece like this one*** [show a spare 3x2 orange piece and put on the table in front of the participant]
- ***2 pieces like this one*** [show a spare 4x2 white piece and put on the table in front of the participant]
- ***3 pieces like this one*** [show a spare 2x2 brown piece and put on the table in front of the participant]

I want to be sure you know what to find because once we start I won't be able to remind you. Can you remember what I asked you to look for?

The participant must demonstrate that they understand which pieces they are searching for. They can do this by referring to the example pieces on the table (e.g. 3 of this one) or by describing the length, width and colour of the pieces. If the participant gets any of the pieces wrong repeat the list in the following way until they are able to remember them correctly (**up to three times only – discontinue if they are unable to learn the list**):

We need:

- ***A piece like this one*** [show the spare 3x2 orange piece and put it back on the table in front of the participant]
- ***Two pieces like this one*** [show the spare 4x2 white piece and put it back on the table in front of the participant]
- ***Three pieces like this one*** [show the spare 2x2 brown piece and put it back on the table in front of the participant]

Then say:

Once you find each piece please put it in this cup [show participant the cup].

There is also something else you need to do during this game. There are some animals like this one (show them a spare example) hidden in the bowls. For each bowl, choose your favourite animal and put it by this cup.

Okay, can you tell me what you need to do in this game?

Ensure child knows: (1) they have to look for specific Lego pieces; (2) these are the pieces that are in front of them on the table (**note: participant does not need to repeat their description/list of these pieces**); (3) they should put the pieces in the cup when they find them; (4) they also need to identify a favourite animal from each bowl; (5) and place these by the cup.

If any of these elements are missing from the child's answer, explain that element to the child.

Next remove the target Lego examples from view.

Allow the participant to continue until they tell you they have finished the task. If they ask for a reminder or help, say:

I can't give you a reminder. Just do your best and let me know when you've finished.

Record the contents of the cup, whether they have selected animals correctly, and the time taken.

⇒ **Discontinue rule:** Discontinue task if child is unable to grasp what is required after all prompts given or fails to start task within 60 seconds of all the instructions being given.

Consent Form Test Part 2

At the end of session, say:

OK – that is the end of our session today [pause to give child time to respond, or wait for 5 seconds if the child does not say anything].

If the participant remembers the consent form now, award 3 points. If they do not remember spontaneously:

Is there anything you want to ask me for before I/you leave?

If the participant remembers the consent form now, award 2 points. If the participant does not ask for the paper, pick up the folder containing the paper and say:

OK I'm going to put this away now – is there anything you want to ask me for?

If the participant remembers the consent form now, award 1 point. If the participant does not ask for the paper, say:

Do you remember you were going to ask me for a piece of paper? Can you show me where it is? (award 0 points regardless of the participant's response)

Appendix VI

Trainee contribution to the project

This project was conducted jointly with another Trainee Clinical Psychologist at University College London (UCL). Having generated a number of ideas and very rudimentary task outlines as a pair, we took ownership of a few of these ideas each. We then worked independently; taking these ideas forward, discussing them with our internal (Dr. Will Mandy) and external supervisors (Professor Paul Burgess), and eventually creating the final Ecologically-Valid Test of Executive Dysfunction (Eco-TED) tasks that were based on them. The Eco-TED was therefore divided in two in order to create two separate empirical papers, which were then written up completely independently. However, we attended research meetings together with our supervisors and contributed to each other's work through discussion in this context.

Data for both studies was collected from the same participants in the same testing session. These testing sessions were carried out by one of the two trainees conducting the research, or one of three trained research assistants.

Appendix VII

Strengths and Difficulties Questionnaire (SDQ)

Strengths and Difficulties Questionnaire

P 4-17

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain or the item seems daft! Please give your answers on the basis of the child's behaviour over the last six months.

Child's Name

Male/Female

Date of Birth.....

	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Restless, overactive, cannot stay still for long	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often complains of headaches, stomach-aches or sickness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Shares readily with other children (treats, toys, pencils etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often has temper tantrums or hot tempers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rather solitary, tends to play alone	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generally obedient, usually does what adults request	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Many worries, often seems worried	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Helpful if someone is hurt, upset or feeling ill	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Constantly fidgeting or squirming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Has at least one good friend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often fights with other children or bullies them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often unhappy, down-hearted or tearful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generally liked by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Easily distracted, concentration wanders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nervous or clingy in new situations, easily loses confidence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kind to younger children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often lies or cheats	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Picked on or bullied by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often volunteers to help others (parents, teachers, other children)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thinks things out before acting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Steals from home, school or elsewhere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gets on better with adults than with other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Many fears, easily scared	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sees tasks through to the end, good attention span	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Do you have any other comments or concerns?

Please turn over - there are a few more questions on the other side

Overall, do you think that your child has difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people?

	No	Yes- minor difficulties	Yes- definite difficulties	Yes- severe difficulties
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

If you have answered "Yes", please answer the following questions about these difficulties:

• How long have these difficulties been present?

	Less than a month	1-5 months	6-12 months	Over a year
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

• Do the difficulties upset or distress your child?

	Not at all	Only a little	Quite a lot	A great deal
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

• Do the difficulties interfere with your child's everyday life in the following areas?

	Not at all	Only a little	Quite a lot	A great deal
HOME LIFE	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
FRIENDSHIPS	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CLASSROOM LEARNING	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
LEISURE ACTIVITIES	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

• Do the difficulties put a burden on you or the family as a whole?

	Not at all	Only a little	Quite a lot	A great deal
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Signature Date

Mother/Father/Other (please specify:)

Thank you very much for your help

© Robert Goodman, 2005

Appendix VIII

School Bag Task individual items analysis results

Item	Measure	<i>N</i>	Control mean (sd)	ASD* mean (sd)	<i>p</i>	Effect size (Cohen's <i>d</i>)
1	Correct items	40	0.95 (0.22)	1.00 (0.00)	0.799	-0.32
1	Additional or incorrect items	40	0.90 (1.02)	2.20 (3.58)	0.242	-0.49
1	Time taken	40	37.75 (31.52)	51.72 (44.98)	0.461	-0.36
1	Time paused	40	0.40 (0.60)	0.55 (0.76)	0.659	-0.22
2	Correct items	40	1.00 (0.00)	1.00 (0.00)	1.000	-
2	Additional or incorrect items	40	0.20 (0.52)	1.60 (3.09)	0.86	-0.63
2	Time taken	40	17.72 (13.19)	32.53 (29.50)	0.149	-0.65
2	Time paused	40	0.75 (0.85)	0.60 (0.82)	0.602	0.18
3	Correct items	40	1.35 (0.88)	1.70 (0.66)	0.265	-0.45
3	Additional or incorrect items	40	0.95 (1.15)	1.10 (1.52)	0.947	-0.11
3	Time taken	40	21.39 (11.05)	32.19 (24.62)	0.201	-0.57
3	Time paused	40	0.65 (0.59)	0.60 (0.75)	0.659	0.07
4	Correct items	40	1.65 (0.49)	1.40 (0.60)	0.242	0.46
4	Additional or incorrect items	40	0.70 (0.87)	1.10 (1.80)	0.620	-0.28
4	Time taken	40	32.12 (22.41)	39.25 (22.79)	0.204	-0.32
4	Time paused	40	0.80 (0.61)	0.55 (0.83)	0.192	0.34
5	Correct items	40	3.00 (0.80)	3.40 (0.68)	0.134	-0.54
5	Additional or incorrect items	40	0.25 (0.44)	0.55 (1.40)	0.904	-0.29
5	Time taken	40	30.81 (20.99)	35.13 (18.53)	0.265	-0.22
5	Time paused	40	0.75 (0.72)	0.40 (0.68)	0.127	0.50
6	Correct items	40	3.15 (0.88)	3.45 (0.83)	0.341	-0.35
6	Additional or incorrect items	40	0.35 (0.49)	1.30 (1.87)	0.174	-0.69
6	Time taken	39	24.60 (11.55)	45.21 (24.11)	0.000	1.09
6	Time paused	40	0.60 (0.68)	0.40 (0.60)	0.398	0.31
7	Correct items	40	2.00 (0.56)	2.15 (0.81)	0.495	-0.22
7	Additional or incorrect items	40	0.90 (0.91)	1.35 (1.63)	0.583	-0.34
7	Time taken	40	33.37 (19.27)	32.06 (23.46)	0.678	0.06
7	Time paused	40	0.90 (0.79)	0.50 (0.69)	0.127	0.54

**Autism Spectrum Disorder*