# Developing and piloting an ecologically valid measure of executive function for children with autism: a function-led approach

Jodie Pullinger

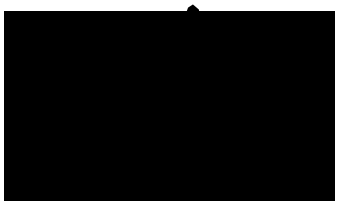
D.Clin.Psy Thesis (Volume 1), 2017

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: Jodie Michele Pullinger

Date: 13th April 2017

# Overview

The focus of this thesis is executive functioning in children and adolescents with Autistic Spectrum Disorders (this term is used synonymously with autism). Part one of the thesis is a systematic review of studies investigating generativity skills using fluency tasks. Specifically, it uses meta-analytic techniques to appraise whether a generativity deficit is evident across the lifespan in those with autism. It then considers variables that may moderate these effects, including participant characteristics and study quality.

The empirical paper (part two) describes the development of a new measure of executive function for children with autism; the EcoTED (Ecologically valid Test of Executive Dysfunction). This measure consists of seven tasks developed using a function-led approach, with the aim of improving on the ecological validity of those measures currently available. The paper reports on the development of four tasks including initial piloting and analysis of their psychometric properties. The project was conducted jointly with another DClinPsy doctoral student who describes the three remaining tasks elsewhere (Bristow, 2016).

The final part of this thesis is a critical reflection on the process of conducting the research. It discusses the origins of the study and the complexities of developing an ecologically valid measure that is psychometrically sound. It gives some suggestions relating to future directions of the Eco-TED and reflects on some of the complexities of research involving those with ASD.

# Table of Contents

## List of Tables

### Part I



### Part II

## List of Figures

### Part 1



### Part 2

## Acknowledgments

First and foremost, thank you to all those that took part in this research. I feel privileged to be able to work with children and families affected by ASD and am humbled by their willingness to give up time and energy to engage in research. I would also like to thank those at Great Ormond Street Hospital, especially Rob, who helped recruit into the study with great positivity.

Next, thanks go to the research team. I am very grateful to my supervisors, Dr Will Mandy and Professor Paul Burgess who gave considerable time and support to this project. Your experience and guidance were invaluable and your passion for research was instrumental in keeping me motivated. Thank you to Eleonora, Aurelie and Salma for your support and contributions to this study. But the biggest thank you must go to Ellie who I really could not have done this without. You have been far more than my research partner – a motivator, a friend, a confidante, who kept me going.

Finally, I would like to thank my family and friends especially my husband and my parents. You have patiently supported me through years of study and had faith that I could achieve the things that I set out to. Even when life threw me a huge curveball and I wasn't sure if I would be able to see this thesis through, you were there with words of encouragement, optimism and boundless love.

# Part I: Literature Review

## A Meta-Analytic Review of Generativity in Autistic Spectrum Disorders

# 1 Abstract

***Aims:*** Children and adults with ASD display executive function deficits, although the precise areas of impairments have not been definitively delineated. One potential aspect of executive function thought to be impaired is generativity (the ability to generate novel responses), although findings have been mixed. This paper offers a quantitative review and synthesis of generativity impairment in ASD.

***Method:*** A systematic search of the literature was carried out using EMBASE, PsycInfo, PubMed and ISI Web of Science. A total of 27 studies featuring child samples and 10 studies featuring adult samples met the inclusion criteria for analysis. These studies utilised mental fluency tasks assessing one or more area of verbal or non-verbal generativity. Verbal Fluency tasks included those using letter exemplars (phonemic fluency tasks) and category exemplars (semantic fluency tasks). Tasks of non-verbal generativity included design and ideational fluency tasks. A weighted, average effect size was calculated for each task using a random-effects model. Moderator variable analysis was carried out based on participant characteristics (IQ and gender) and study quality.

***Results:*** The findings of this review support an overall impairment on fluency tasks for both children and adults with ASD compared with typically developing controls. Effect sizes varied from small to moderate with the largest impairment on tasks assessing phonemic fluency (child studies: ES = -0.82, 95% CI = -1.08 to -0.57; adult studies: ES = -0.59, 95% CI = -0.87 to -0.31). The smallest effect size was noted for design fluency tasks in children (ES = -0.28, 95% CI = -0.54 to -0.03). There was substantial between-study variability for phonemic fluency tasks ($I^2$=59.9%) and semantic tasks in children ($I^2$=72.9%). Categorical analysis of moderator variables

suggested that effects may be larger in studies rated as poorer quality and for ASD participants with lower IQ.

***Conclusion:*** This review offers support for a generativity impairment in children and adults with ASD. The large degree of heterogeneity across studies may be a consequence of study quality or may reflect the difficulties in comparing executive function abilities using current measures.

# 2 Introduction

## 2.1 Autistic Spectrum Disorder

Autistic Spectrum Disorder (ASD) is a neurodevelopmental condition characterised by pervasive patterns of social communication deficits along with repeated, repetitive patterns of behaviour or interests (American Psychiatric Association, 2013). It is estimated to affect around 1.1% of the population (Fombonne, Quirke, & Hagen, 2011) with a marked difference in prevalence of diagnosis amongst males compared with females (Baird, 2006). However, given increasing diagnosis rates of the disorder this is believed to be a conservative estimate, particularly as it is widely recognised that ASD is under-diagnosed in females (Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2015). There are no known distinct biological markers for ASD and so it is a diagnosis that relies on observed behavioural features. Those with ASD show differences in a range of areas including social, emotional, motor and cognitive functioning. As such, impairments in individuals with ASD can range from mild to severe and may be associated with increased co-morbidity (43.9%) of a learning disability (Fombonne et al., 2011).

## 2.2 Executive Function

The range of abilities displayed by those with ASD are associated with aspects of executive control. Anderson (1998) describes executive functions as those skills which are "necessary for purposeful goal-directed behaviour". They include an array of cognitive processes including working memory, response inhibition, planning and set shifting. It has been posited that deficits in these areas, may be key to understanding behavioural phenotypes of those with ASD (the executive dysfunction hypothesis) (Ozonoff, Pennington, Bruce, & Rogers, Sally, 1991). These include cognitive inflexibility, restricted interests and adherence to routines (Hill, 2004a).

An influential evaluation of the executive dysfunction hypothesis was conducted by Hill (2004a) who, in a non-systematic, narrative review, highlighted a range of frontal lobe processes that may be affected in those with ASD. Since this article there have been a number of published systematic quantitative reviews focusing on executive domains including; flexibility, memory, visuo-spatial performance and inhibition (Bordignon, Endres, Trentini, & Bosa, 2015; Hill, 2004a; Leung & Zakzanis, 2014b; Muth, Hönekopp, & Falter, 2014). However, one key area of impairment that Hill (2004b) identified was generativity and this has not been investigated in a published, systematic, quantitative review.

## 2.3 Generativity

Generativity is a cognitive activity that is perhaps less well conceptualised than most other elements of executive function (Dichter, Lam, Turner-Brown, Holtzclaw, & Bodfish, 2009). It is described as the ability to spontaneously generate novel ideas without excessive pauses or errors (Pastor-Cerezuela, Fernandez-Andres, Feo-

Alvarez, & Gonzalez-Sala, 2016; Turner, 1999). The term is often used synonymously with mental fluency (e.g. Bishop & Norbury, 2005a). The ability to generate novel responses can be observed across both verbal and non-verbal tasks and as such a range of different types of generativity have been proposed including verbal, ideational and design fluency. It has also been linked with more abstract concepts such as creativity and imagination (Craig & Baron-Cohen, 1999). However, whilst tests of creativity and imagination tend to place more emphasis on the quality and value of ideas generated, e.g. as in the Torrance Test of Creative Thinking (Torrance, 1974), tests of generativity are primarily concerned with the number of novel responses produced. Generativity impairments have been reported in a number of neuropsychiatric and neuropsychological disorders (Vannorsdall, Maroof, Gordon, & Schretlen, 2012).

## 2.4 Measures of Generativity

Generativity can be broadly separated into two domains; verbal and non-verbal. Within each of these domains there are different forms of generativity that are measured using a variety of means. Verbal generativity is commonly assessed using verbal fluency tasks which are short tests of verbal functioning (see Lezak, Howieson, Bigler, & Tranel, 2012). In these tasks individuals are provided with either phonemic (letter) or semantic (category) cues and asked to generate as many words as possible within a given time (usually 60 seconds). These tasks require the participant to access information stored in their lexicon.

Design fluency tasks were created as a non-verbal analogue to the above. They most commonly feature in neuropsychological batteries including the Developmental NEuroPSYchological Assessment (NEPSY)(Korkman, Kirk, & Kemp, 1998) and the

Delis-Kaplan Executive Function System (D-KEFS™) (Delis, Kaplan, & Kramer, 2001). Participants are given an array of structured and unstructured dots and asked to create as many novel designs as possible by joining the dots using a set number of straight lines. Unlike verbal tasks, design fluency tasks do not rely on stored knowledge (Turner, 1999b).

Ideational fluency can be conceptualised as a further sub-domain of non-verbal generativity and is thought to measure divergent thinking; or the ability to generate creative ideas and de novo responses (Snyder, Mitchell, Bossomaier, & Pallier, 2004; Turner, 1999a). Some authors use the term "ideational fluency" to refer more generally to the ability to "produce a large number of ideas" (Hocevar, 1979, p.191) and therefore include verbal and design fluency abilities under this umbrella term (Vannorsdall et al., 2012). However, more commonly it is used to describe non-verbal generativity abilities that are assessed in one of two ways. In variants of the 'uses of objects task', participants are shown various items and are required to come up with as many novel uses for an object as possible in a set time frame (typically 150 seconds). Examples of those types of objects used include a brick, pencil and a mug. The second type of task used to assess ideational fluency is the Pattern Meanings Task (Wallach & Kogan, 1965). In this task individuals are shown a range of meaningless line drawings and asked to generate ideas of what the drawings could represent. See Table 1.1 for a more thorough summary of tasks.

## 2.5 Generativity and ASD

In a seminal paper by Turner (1999) individuals with autism demonstrated significantly fewer imaginative responses on word and ideational fluency tasks when

compared to a control group. The author concluded that this was evidence for a generativity deficit hypothesis; and proposed this as the foundation for the restricted behaviour and lack of spontaneity observed in children with ASD. Other studies have replicated these findings (Barron-Linnankoski et al., 2015; Begeer et al., 2014; Dichter et al., 2009) whilst also providing evidence for poorer performance on design fluency tasks (Narzisi, Muratori, Calderoni, Fabbro, & Urgesi, 2013). Rutherford & Rogers (2003) found a strong association between impaired generativity and lack of pretend play whilst others have also linked it to impoverished imagination in children with ASD (Craig & Baron-Cohen, 1999; Jarrold, Boucher, & Smith, 1996). In addition, functional imaging studies have demonstrated a distinct profile of frontal lobe activation in those with ASD undergoing fluency tasks, lending support to the idea that distinct neurocognitive processes underlie generativity and that these differ in children with and without the disorder (Frith, Friston, Liddle, & Frackowiak, 1991; Gaillard et al., 2000).

Table 1.1 Measures of Generativity

| Task / Measure | Brief Description |
|---|---|
| Verbal Fluency Task / Word Fluency Task (Strauss et al 2006)[a] | Participants are asked to produce spontaneous novel responses (words) within a set period (e.g. to name as many words as possible beginning with 's' in one minute). |
| Phonological / Letter Fluency[a] | Generation of words that start with a specific letter within a given time frame. |
| Controlled Oral Word Association Test[a] (COWAT) (A. Benton & Hamsher, 1983) | Generate as many words as possible in one minute beginning with the letters 'F', 'A' and 'S'. Participants are instructed that responses should not be proper nouns or repeated words with different endings (e.g. fly and flying). |
| Chicago Word Fluency Test[a] | Participants must write as many words as possible that begin with a specific letter within a set period. |
| D-KEFS Verbal Fluency Test (Delis et al., 2001)[ab] | Measures letter fluency, category fluency and category switching. Participants are asked to provide as many words as possible beginning with the letters 'F', 'A' and 'S'.<br>For category fluency, they are required to produce as many responses as possible for the categories of 'animals' and 'boy's names'. In the switching condition, they have to produce as many responses as possible for the categories of 'fruits' and 'furniture' whilst switching between the two. There is a time limit of one minute applied to each condition. The authors state that this task is a measure of fluent productivity in the verbal domain. |
| Animal Fluency Task (M. Lezak, 1995)[b] | Participants are asked to name as many animals as possible within a set time. |
| Use-of-objects Task (Turner, 1999b)[c] | Participants are shown six objects one at a time and asked to generate as many uses for them as possible. Half of the objects have obvious conventional uses whilst half do not. |

| | |
|---|---|
| | Those objects routinely used are; brick, pencil, mug, 3' dowel rod, 3' piece of ribbon/fabric, 32" piece of elastic. An example is provided for each object. In the case of the conventional objects this includes one conventional and one novel use. The participant is then asked, 'tell me other ways in which you think this object could be useful'. A prompt is given if the participant does not provide a response for 15 seconds; 'keep thinking, how else could it be useful?'. The time given for this task is routinely 2.5 minutes but differs between studies. |
| Ideational Fluency (Turner, 1999b)[c] | Generating novel uses for a specific object (e.g. how many uses can be found for a hat) within a given time frame. An example of objects includes: newspaper, brick, pencil, mug and toothpick |
| Pattern Meanings (Wallach & Kogan, 1965)[c] | Participants are shown a meaningless line drawing on a card and asked, 'what could it be?' Examples are provided by the examiner. They are then shown 5 more meaningless line drawings and asked to come up with different responses as to what the drawings could represent in a given time (usually 1.5 minutes). The participant can change the orientation of the card during the task. |
| D-KEFS Design Fluency Test (Delis et al., 2001)[d] | Participants are provided with grids containing an array of dots and asked to produce as many different patterns as they can by connecting the dots using four straight lines only. Each line must touch at least one other line at a dot. There are three conditions for this task; (1) grids contain five filled black dots and designs must connect the filled dots; (2) grids contain five filled and five empty dots and participants must connect the empty dots only; (3) a |

| | |
|---|---|
| | switch condition where participants asked to draw designs by switching back and forth between empty and filled dots. |
| Design Fluency Test (Jones-Gotman & Milner, 1977)[d] | There are two parts to this task. First, participants are instructed to produce as many novel drawings as possible that do not represent actual objects or nameable patterns (such as geometric shapes). They are provided with two examples of drawings that would be acceptable. They are then given five minutes to produce as many drawings as possible. One warning is given should the participant violate the rules by drawing a real or nameable object or a design that is too like another. In the second part of the task participants are given four minutes to produce as many drawings as possible that consist of exactly four lines. These could be straight, curved or circular in nature. |

a. *These tasks all measure verbal or phonemic fluency although the demands of the task may vary slightly in terms of the letters that are used, time allowed and mode of response.*
b. *These tasks measure semantic (category) fluency although there may be slight variations in administration and the category cues used.*
c. *These tasks provide a measure of ideational fluency.*
d. *These are a measure of design fluency (fluent productivity in the spatial/non-verbal domain* (Delis et al., 2001).

Despite this evidence there have also been a number of studies that have found no significant differences between those with ASD and those without (Goddard, Dritschel, Robinson, & Howlin, 2014; Robinson, Goddard, Dritschel, Wisley, & Howlin, 2009). There are several reasons why this may be the case. Firstly, inconsistent findings may reflect the degree of heterogeneity amongst those with ASD.

For example, functional imaging studies suggest that neural connectivity varies markedly in those with ASD; and that some on the less severe end of the spectrum may be neuro-anatomically similar to their TD counterparts (Lenroot & Yeung, 2013; Uddin, 2015). Therefore, studies that include mainly high-functioning participants with low symptom severity, may be less likely to show generativity impairments than studies including participants with marked ASD symptomology.

Secondly, in her non-systematic, qualitative synthesis of the literature, Hill (2004) made the point that studies often fail to ensure that participants are adequately matched. They have included clinical and control groups with significantly different IQ, age and gender ratios (Barron-Linnankoski et al., 2015; Happe, Fleminger, & David, 2006). These have all been shown to effect fluency abilities (Ardila, Pineda, & Rosselli, 2000; Spek, Schatorje, & Scholte, 2009). Differences in findings across studies may therefore be a consequence of sample characteristics.

It is also important to note that studies use a variety of different "hybrid" tasks to measure generativity, thus limiting their between-study comparability (Shao, Janse, Visser, & Meyer, 2014). Despite evidence of decreased generativity performance as a function of time (Crowe, 1998), tasks have varied with regards to the amount of time allowed for the participant to respond (e.g. see Bishop & Norbury, 2005b; Turner, 1999a). There are also a variety of objects used for ideational fluency tasks and differing category exemplars for verbal tasks. Therefore, uncertainties in the literature may be further influenced by changes in task administration.

Finally, there is the criticism that many studies within the generativity literature lack statistical power due to small sample size. This makes the area of study more susceptible to reification of sampling error and high rates of type II error (Henry &

Crawford, 2004). Meta-analytic techniques and the assimilation of studies is one way to reduce bias, increase statistical power, and investigate which of these additional factors might moderate findings (Field & Gillett, 2010; Petticrew & Gilbody, 2004; Sharpe, 1997).

## 2.6 Rationale for Present Quantitative Review

To our knowledge there are currently no quantitative systematic reviews of generativity in individuals with ASD compared to TD controls. This is surprising given that generativity has been posited to underlie a range of important cognitive processes, symptomology and functional impairments. Studies that have looked at generativity in both adult and child populations have reported mixed findings which has led to confusion about the role generativity plays in observed behaviours. Not only have these studies reported mixed findings in terms of significance, there have also been studies that have suggested slightly superior performance in those with ASD (e.g. Lind & Bowler, 2010). Articles that have reviewed the literature have been narrative in focus, have failed to draw clear conclusions and are now out of date (Hill, 2004a). Furthermore, generativity is a cognitive process that lends itself to meta-analytic review. It is measured using only minor variations of the same fluency tasks; and that it is hypothesised to be a relatively discrete neurocognitive process (Dichter et al., 2009).

Therefore, the aims of this paper are to provide a comprehensive quantitative review of generativity in individuals with ASD. More specifically it aims to:

- Discover whether those with ASD are impaired on tasks of generativity compared to TD controls

- Investigate whether generativity deficits are evident for both verbal and non-verbal fluency tasks
- Considers whether factors such as gender, IQ and study quality influence findings on generativity in ASD compared to TD controls

# 3 Method

## 3.1 Search Strategy and Selection Criteria

Key concepts were isolated and a list of synonyms were created to ensure that the search incorporated all possible definitions of generativity and autism. These were created through identifying subject headings and keywords tagged within major articles (Hill, 2004b; Kenworthy, Yerys, Anthony, & Wallace, 2008). A systematic search of EMBASE, PsycINFO, PubMed and ISI Web of Science Cross Search (Thomson Scientific/Institute for Scientific Information Web Services) was conducted on 20th, 22nd and 27th August 2016. By monitoring the yield of included studies from these databases it was apparent that a search of these, along with the reference lists of key papers, was sufficient (Field, 2000; Petticrew & Gilbody, 2004).

Preliminary search terms were [autism OR pervasive developmental disorders OR Asperger OR Asperger syndrome OR ASD OR autis* OR autism spectrum disorders OR autistic OR autistic thinking OR high functioning autism OR social communication disorder OR PDD OR PDD-NOS] AND [generativity]. From these specific measures of interest were identified which were then combined with the autism search terms [verbal fluency OR ideational fluency OR design fluency OR semantic verbal fluency OR word fluency OR category fluency OR letter fluency OR use of objects task OR animal fluency task OR COWAT OR DKEFS OR NEPSY].

Duplicates were removed following the preliminary search and the remaining articles were assessed per the inclusion and exclusion criteria outlined below to avoid selection bias (Rosenthal, 1990). This methodology follows that suggested by the PRISMA Statement for reporting meta-analysis (Moher, Liberati, Tetzlaff, Altman, & GROUP, 2009) (Figure 1.1).

### 3.2 Inclusion Criteria

Studies were included in the analyses if they met the following inclusion criteria: (1) empirical articles that assessed generativity; (2) included one or more variants of a fluency task, regardless of whether the primary goal of the paper was an assessment of generativity; (3) participants were either children and adolescents up to the age of 18 years or adults over the age of 18 years; (4) ASD participants satisfied the formal diagnostic criteria according to the Diagnostic and Statistical Manual of Mental Disorders (3rd, 4th or 5th editions) (American Psychiatric Association, 1987, 2000, 2013) and/or the International Statistical Classification of Diseases and Related Health Problems, 10th revision (World Health Organisation, 2012); (5) included a comparison group of typically developing controls; (6) were published between 1980 and 2016 in a peer reviewed journal.

### 3.3 Exclusion Criteria

Studies were excluded from the analyses if: (1) they were not published in English; (2) were single case studies or case reports; (3) participants were given treatment in the form of medication or an intervention that might have altered executive function performance; (4) samples included individuals with Fragile X

syndrome, traumatic brain injury or other neurological disorders; (5) they were brain imaging studies that did not include a generativity task performed outside of scanning conditions; (6) they specifically targeted language and/or reading fluency rather than a measure of generativity; (7) included a mixed sample of children, adolescents and adults unless the data were presented separately for each. Papers were also excluded if they did not contain sufficient information to allow for statistical analyses even after contact with the authors had been sought.

### 3.4 Studies Included in Analysis

Providing studies met the criteria, they were included in the analysis regardless of whether generativity was the focus of their investigation. This is because several studies looking at other areas of executive function, such as autobiographical memory, also included fluency tasks.

Two studies reported data in graphical form only (Mostert-Kerckhoffs, Staal, Houben, & Jonge, 2015; Turner, 1999a). In these instances, means were estimated from the graphs and standard deviations were calculated using confidence intervals according to the following method (Higgins & Green, 2008):

$$SD = \sqrt{N}\ (upper\ limit - lower\ limit)/3.92$$

These studies were identified as key papers within the literature and so it was felt that they were fundamental to the review. Authors of both papers were approached for precise parameter estimates but failed to reply. Obtaining approximate effect size estimates using the data which is available is not unusual (Rosenthal, 95). However, to ensure that the accuracy of these methods did not adversely affect the overall analyses, analyses were run with and without these datasets to ensure that they did not

skew the results. Findings with and without the inclusion of these studies were substantively the same, therefore they were included in the final analysis.

Where studies split their samples into groups based on higher or lower IQ or per their ASD diagnosis, the groups were combined to create just one clinical and one control group. Means and standard deviations were calculated using the methods suggested in Higgins & Green (2008).

### 3.5 Choice of Measures

Specific measures were chosen based on the commonly cited seminal paper by Turner (1999a) which outlined tasks assessing verbal and non-verbal generativity. Preliminary searches indicated that empirical studies investigating verbal generativity all included a measure of phonemic and/or semantic generativity that were a version of that used in the Turner paper. The tasks had only minor variations in administration or letter / category exemplars (as demonstrated in Table 1.1). For this reason we included all measures that were a variation of the original 'FAS' verbal fluency task featured in the Turner paper (Benton, 1968). For non-verbal generativity, Turner (1999a) identified two sub-domains; ideational and design fluency. Tasks use to measure the latter were again largely the same across the literature with only slight differences in administration and were universally referred to as "design fluency tasks". These were therefore included in the current review. Pattern Meanings and the Use of Objects Task were chosen as analogous measures of ideational generativity as they have been shown to be highly correlated; are thought to measure the same cognitive processes; and are consistently used in the literature as measures of ideational fluency (Bereiter, Harris, Archer, & Klausmeier, 1960; Chan et al., 2001;

Wallach & Kogan, 1965). Synthesising data in this way is not unfamiliar and has been done for fluency tasks as well as other executive functions (e.g. Henry & Crawford, 2004; Leung & Zakzanis, 2014b).

### 3.6 Dependent Variables of Measures

There was some variation in the dependent variables reported for the generativity tasks across studies. Those studies that evaluated verbal (semantic and phonemic) generativity predominantly recorded the total mean number of correct words. Others reported the mean number of correct words per category cue, the proportion of correct words, or scaled scores of overall performances. Of the 31 studies included in the final analysis of verbal generativity, all but two reported the mean number of correct scores (across the task or per category cue). This was therefore chosen as the variable for analysis. When this was not available, the proportion of correct responses (i.e. Low, Goddard and Melzer, 2009) or the scaled scores (i.e. Narzisi et al., 2013) were selected as the variable that were most closely related. Two studies reported a verbal composite of the combined semantic and phonemic performance and so were not included in the final analysis (Barron-Linnankoski et al., 2015; Koolen, Vissers, Egger, & Verhoeven, 2014). To reduce bias, dependent variables were selected before the data was extracted.

For design fluency, each of the four studies included in the analysis recorded the total number of correct patterns as their dependent variable so these were directly compared. For ideational fluency, which was measured through variants of the 'use of Objects' tasks, again the dependent variable differed across the six included studies. Three of the studies reported the total number of possible uses across all the cues; two

of the studies reported the average number of uses per object, and one study reported the proportion of correct responses. Combining dependent variables that measure the same underlying concept (in this case generativity) is not uncommon in meta-analyses (Wykes, Huddy, Cellard, McGurk, & Czobor, 2011).

# 4 Statistical Procedures

## 4.1 Effect Size Calculation

Multiple meta-analytic procedures were employed using STATA 12 (StataCorp LP, 2011). Firstly, estimation of a summary statistic for each of the studies in the form of Cohens *d* was calculated. A negative Cohens *d* statistic represented poorer performance of the autistic individuals. All effect sizes were calculated using means and standard deviations.

This was followed by weighted average of effect sizes across studies to estimate the 'true effect size' in the population (Deeks, Altman, & Bradburn, 2008; Field, 2000). In instances where there was significant heterogeneity, further exploration of moderator variables was conducted.

## 4.2 Heterogeneity of Effect Sizes

Heterogeneity of effect size was quantified using Cochrane's Q and the $I^2$ statistic (Harris et al., 2008; Hedges & Olkin, 1985; Higgins & Thompson, 2002). Heterogeneity may be caused by variability in sampling error (a consequence of using different samples) or between studies variability. The latter represents true heterogeneity between effect sizes due to moderating factors such as sample characteristics, study quality and variations in the fluency tasks (Huedo-Medina,

Sánchez-Meca, Marín-Martínez, & Botella, 2006). In using the Q and $I^2$ statistic, both sources of variance were accounted for (Higgins & Thompson, 2002). Further, Q alone can lack power to detect heterogeneity when based on a small number of studies, which was particularly the case for the non-verbal fluency tasks (Huedo-Medina et al., 2006).

A random-effects model using the DerSimonian and Laird method was chosen for analyses (Deeks et al., 2008). This model incorporates an estimate of heterogeneity in the weighting and therefore allows for inferences to be generalised beyond the study (Field & Gillett, 2010; Harris et al., 2008; Hedges, 1992).

### 4.3 Publication Bias

As the analyses included studies published in peer reviewed journals only, tests for potential publication bias were included to address the 'file drawer effect' (Rosenthal & Rubin, 1979; Sharpe, 1997). This was assessed visually by way of a funnel plot and statistically using the test proposed by Egger (Egger, Davey Smith, Schneider, & Minder, 1997). The latter is specifically suggested for examining asymmetry in continuous outcomes (Higgins & Green, 2008). In instances when potential publication bias was detected, Rosenthal's failsafe-N (Rosenthal, 1979) was calculated to see whether the effect size estimate was robust, i.e. how many unpublished studies would be required to overturn calculated mean effect sizes (Ellis, 2010).

## 5 Results

A flow diagram depicting the number of papers retrieved and included in the final analysis is shown in Figure 1.1. A summary of the studies included in the final analyses are displayed in Tables 1.3 and 1.4 along with calculated effect sizes.

## 5.1 Study Quality

To address validity threats (as discussed in Sharpe, 1997), study quality was assessed for each of the papers included in the final analyses. Existing quality assessment tools such as the CASP Case Control Study Checklist (Critical Appraisal Skills Programme (CASP), 2014) were considered. However, there was limited operational utility when applied to the current review as studies were not interventional/treatment studies.

Criteria used to assess quality were therefore adapted from the Standard Quality Assessment Criteria for Evaluating Primary Research Papers (Kmet, Lee, & Cook, 2004) to include criteria that were specifically relevant to the current analyses. An example of the adapted tool along with specific criteria are included in Appendix 1. Evaluation scores for each of the studies are shown in Table 1.2.

Figure 1.1 Flow diagram depicting search strategy according to PRISMA criteria (Moher et al., 2009)

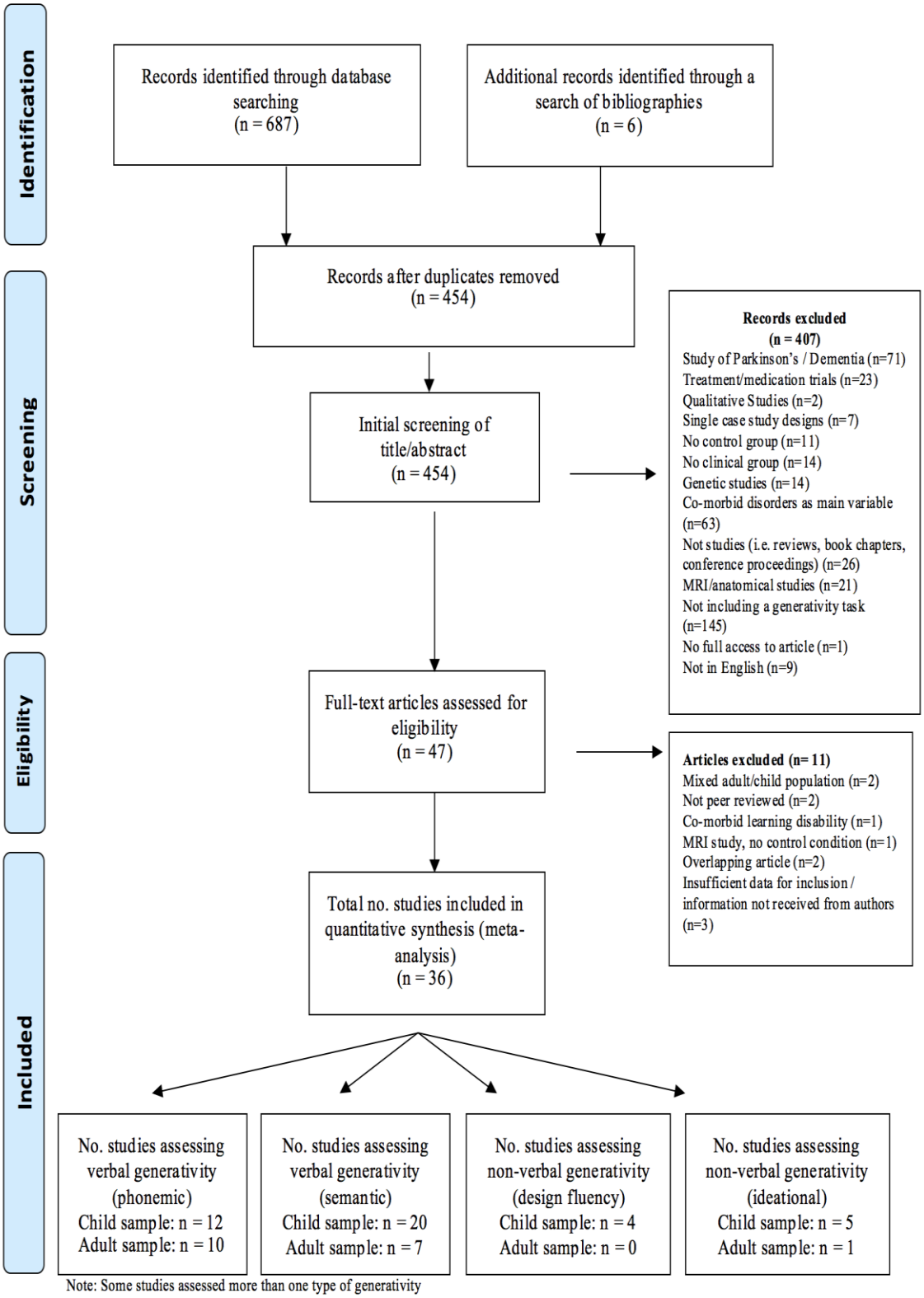


Table 1.2. Study Evaluation Tool

| **Study** | Methodology (/16) | Reporting (/8) | Overall Score (/24) |
|---|---|---|---|
| ***Child Samples*** | | | |
| Van Eylen et al. (2015) | 16 | 8 | 24 |
| Goddard et al. (2014) | 15 | 8 | 23 |
| Robinson et al. (2009) | 14 | 8 | 22 |
| Verté et al. (2005) | 13 | 8 | 21 |
| Geurts et al. (2004) | 13 | 8 | 21 |
| Williams et al. (2006) | 13 | 7 | 20 |
| Pastor-Cerezuela et al. (2016) | 12 | 8 | 20 |
| Dichter et al. (2009) | 12 | 7 | 19 |
| Happé et al. (2006) | 12 | 7 | 19 |
| Goddard and Dritschel (2014) | 11 | 8 | 19 |
| Kilinçaslan et al. (2010) | 11 | 8 | 19 |
| Mostert-Kerckhoffs et al. (2015) | 13 | 5 | 18 |
| Bishop and Norbury (2005) | 10 | 8 | 18 |
| Corbett et al. (2009) | 10 | 8 | 18 |
| Panerai et al. (2014) | 10 | 7 | 17 |
| Craig and Baron-Cohen (1999) | 9 | 8 | 17 |
| Dunn et al. (1996) | 9 | 8 | 17 |
| Low et al. (2009) | 9 | 7 | 16 |
| Barron-Linnankoski et al. (2015) | 9 | 7 | 16 |
| Czermainski et al. (2014) | 9 | 7 | 16 |
| Hanson and Atance (2014) | 9 | 7 | 16 |
| Chan et al. (2011) | 8 | 8 | 16 |
| Mashal and Kasirer (2011) | 8 | 8 | 16 |
| Boucher (2009) | 8 | 7 | 15 |
| Weismüller et al. (2015) | 8 | 7 | 15 |
| Narzisi et al. (2013) | 8 | 6 | 14 |

| | | | |
|---|---|---|---|
| Turner (1999) | 7 | 6 | 14 |
| ***Adult Samples*** | | | |
| Lever and Geurts (2016) | 14 | 8 | 22 |
| Lind and Bowler (2010) | 14 | 8 | 22 |
| Bramham et al. (2009) | 13 | 8 | 21 |
| Spek et al. (2009) | 12 | 7 | 19 |
| Geurts and Vissers (2012) | 12 | 8 | 20 |
| Inokuchi and Kamio (2013) | 11 | 8 | 19 |
| Lopez et al (2005) | 11 | 8 | 19 |
| Ambery et al. (2006) | 11 | 7 | 18 |
| Rumsey and Hamburger (1988) | 10 | 7 | 17 |
| Kleinhans et al. (2008) | 8 | 7 | 15 |

## 5.2 Verbal Generativity

### *5.2.1 Phonemic Fluency Tasks*

Twelve child samples and ten adult samples entered the analysis of verbal generativity as measured by phonemic fluency tasks. Those studies involving children included a total of 326 ASD participants (mean age = 10.65 years, SD = 1.42) and 309 TD controls (mean age = 10.59, SD = 1.06). The ten adult studies yielded a total of 349 participants with ASD (mean age = 36.11 years, SD = 13.01) and 295 TD controls (mean age = 36.17 years, SD = 13.06).

The results of the combined and subgroup analysis of both the child and adult studies for letter fluency tasks are presented in Figure 1.2. The overall effect of ASD on verbal generativity as assessed by letter fluency tasks was estimated to be -0.72 (95% CI = -0.91 - -0.53). This is a medium effect size according to Cohen (1992). The effect size, expressing impairments in ASD compared to controls, was greater in children ($d$ = -0.82, CI = -1.08 - -0.57) than for adults ($d$ = -0.59, CI = -0.87 – -0.31),

although it should be noted that there were several studies showing no effect at the study-level. There was a significant group difference overall ($p<0.001$) and for the child ($p<0.001$) and adult studies ($p<0.001$).

Heterogeneity was significantly greater than zero for both the child ($Q = 25.22$, $df = 11$, $p<0.01$, $I = 56.4\%$) and the adult studies ($Q = 22.28$, $df = 9$, $p<0.01$, $I = 59.6\%$) and was highly significant overall ($Q = 52.30$, $df = 21$, $P<0.001$, $I = 59.9\%$).

**Figure 1.2** Forest plot for studies comparing generativity performance in phonemic fluency tasks between ASD and TD for both child and adult samples

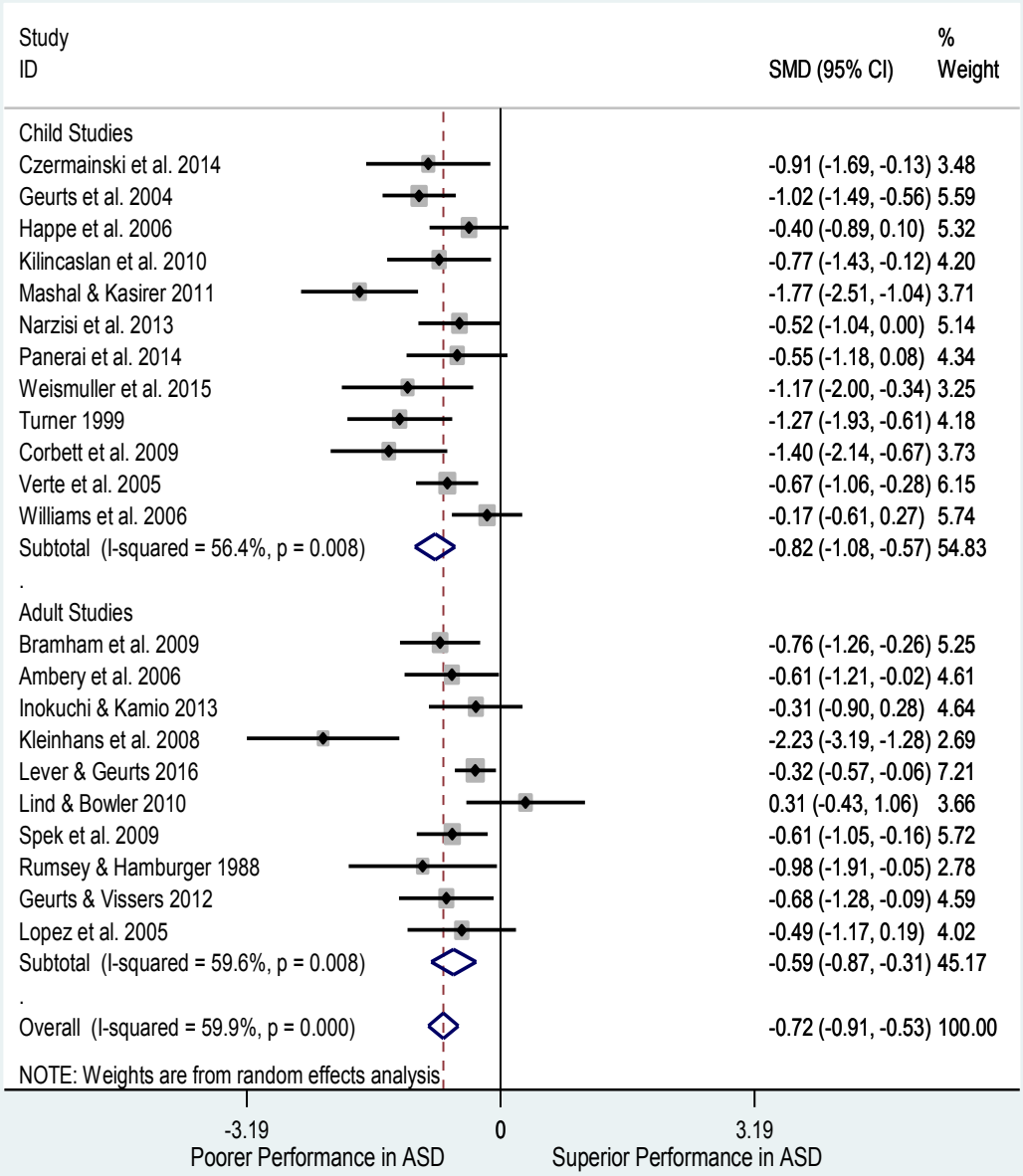


**Table 1.3** Studies Comparing Generativity in Samples of Children and Adolescents with ASD and TD Controls

| **Article** | **Sample** | **Age (years) (M±SD)** | **Gender (% male)** | **IQ (M±SD)** | **Type of Generativity Task** | **Effect Size (d)** |
|---|---|---|---|---|---|---|
| Czermainski et al. (2014) | 11 Autism/Asp | 11.73 (1.6) | NR | NR | Verbal Fluency (Phonemic) | -0.91 |
| | 19 TD | 11.42 (1.8) | | | Verbal Fluency (Semantic) | -0.86 |
| Geurts et al. (2004) | 41 HFA | 9.4 (1.8) | 100 | 98.3 (18.4) | Verbal Fluency (Phonemic) | -1.02 |
| | 41 TD | 9.10 (1.7) | 100 | 111.5 (18) | Verbal Fluency (Semantic) | -0.93 |
| Happé et al. (2006) | 32 AS/HFA | 10.9 (2.4) | 100 | 99.7 (18.7) | Verbal Fluency (Phonemic) | -0.40 |
| | 32 TD | 11.20 (2) | 100 | 106.8 (3.4) | Verbal Fluency (Semantic) | -0.41 |
| | | | | | Non-verbal Fluency (Design) | -0.31 |
| Kilinçaslan et al. (2010) | 21 Asp | 12.44 (2.87) | 85.7 | 105.52 (14.74) | Verbal Fluency (Phonemic) | -0.77 |
| | 18 TD | 11.96 (2.36) | 83.3 | 107.27 (13.39) | Verbal Fluency (Semantic) | -0.48 |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Mashal and Kasirer (2011) | 20 ASD | 13.02 (NR) | 90 | NR | Verbal Fluency (Phonemic) | -1.77 |
| | 20 TD | NR | 90 | NR | Verbal Fluency (Semantic) | -2.13 |
| Narzisi et al. (2013) | 22 HFA | 9.77 (3.65) | 100 | 99.09 (14.23) | Verbal Fluency (Phonemic) | -0.52 |
| | 24 TD | NR | 100 | NR | Verbal Fluency (Semantic) | -0.50 |
| Panerai et al. (2014) | 19 ASD/HFA | 9.23 (3.31) | 78.9 | NR | Verbal Fluency (Phonemic) | -0.55 |
| | 21 TD | 9.73 (2.62) | 73.5 | NR | Verbal Fluency (Semantic) | -0.59 |
| Weismüller et al. (2015)[a] | 15 Autism | 9.4 (2.4) | 100 | 99.3 (18.4) | Verbal Fluency (Phonemic) | -1.17[a] |
| | 12 TD | 10.60 (3.25) | 100 | 118.3 (15.1) | Verbal Fluency (Semantic) | -1.45[a] |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Turner (1999)[b] | 22 HFA | 12 (5.4) | 86.4 | 100 (22.3) | Verbal Fluency (Phonemic) | -1.27[b] |
| | 21 TD | 11.11 (4.5) | 85.7 | 101 (17.8) | Verbal Fluency (Semantic) | -1.33 [b] |
| Corbett et al. (2009) | 18 HFA/Asp | 9.44 (1.96) | 94.4 | 94.17 (17.79) | Verbal Fluency (Phonemic) | -1.40 |
| | 18 TD | 9.56 (1.81) | 66.7 | 112.22 (14.84) | Verbal Fluency (Semantic) | -1.36 |
| Verté et al. (2005) | 61 HFA/Asp | 9.1 (1.9) | 85.07 | 99.7 (17.1) | Verbal Fluency (Phonemic) | -0.67 |
| | 47 TD | 9.40 (1.6) | 85.1 | 112.1 (9.7) | Verbal Fluency (Semantic) | -0.93 |
| Williams et al. (2006) | 44 Autism | 11.36 (2.18) | 82.1 | 104.13 (15.09) | Verbal Fluency (Phonemic) | -0.17 |
| | 36 TD | 11.82 (2.2) | 69.6 | 107.5 (8.21) | | |
| Robinson et al. (2009) | 54 HFA/Asp | 12.54 (2.80) | 77.8 | 103.53 (10.54) | Verbal Fluency (Semantic) | -0.11 |
| | 54 TD | 12.08 (2.34) | 77.8 | 104.8 (9.07) | | |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Dichter et al. (2009) | 39 ASD/Autism | 9.72 (2.66) | 97.4 | 101.69 (17.5) | Verbal Fluency (Semantic) | -0.84 |
| | 39 TD | 10.57 (3.35) | 97.4 | 111.67 (16.11) | Non-verbal Fluency (Ideational) | -0.62 |
| Boucher (2009) | 7 Autism | 14.2 (1.00) | 100 | NR | Verbal Fluency (Semantic) | -0.34 |
| | 7 TD | 13.10 (1.43) | 100 | NR | | |
| Chan et al. (2011) | 16 AS/HFA | 8.00 (1.90) | 100 | 89.5 (18.23) | Verbal Fluency (Semantic) | -0.91 |
| | 19 TD | 8.30 (1.98) | 100 | 101 (20.65) | | |
| Dunn et al. (1996) | 10 HFA | 6.79 (1.9) | NR | 102.4 (10.06) | Verbal Fluency (Semantic) | 0.61 |
| | 10 TD | 4.93 (1.51) | NR | 106.4 (12.1) | | |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Goddard and Dritschel (2014) | 24 ASD | 12.89 (2.08) | 50 | 105.9 (12.8) | Verbal Fluency (Semantic) | 0.09 |
| | 24 TD | 12.57 (2.02) | 50 | 106.3 (10.8) | | |
| Goddard et al. (2014) | 63 ASD | 12.58 (2.81) | 81.0 | 103.6 (20.51) | Verbal Fluency (Semantic) | 0.01 |
| | 63 TD | 12.10 (2.26) | 81.0 | 104.76 (11.79) | | |
| Hanson and Atance (2014) | 25 ASD/Autism | 5.86 (1.49) | 88 | 85.71 (21) | Verbal Fluency (Semantic) | -0.11 |
| | 25 TD | 4.86 (0.93) | 88 | 109.12 (8.03) | | |
| Pastor-Cerezuela et al. (2016) | 47 ASD | 6.67 (1.14) | 85.1 | 98.89 (19.52) | Verbal Fluency (Semantic) | -0.85 |
| | 53 TD | 6.74 (1.08) | 81.1 | 99.64 (16.75) | | |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Barron-Linnankoski et al. (2015) | 30 HFA/Asp | 9.10 (1.3) | 93.3 | 107.2 (17.3) | Non-verbal Fluency (Design) | -0.01 |
| | 60 TD | 9.10 (1.4) | 90 | 104.2 (20.9) | | |
| Bishop and Norbury (2005) | 14 HFA | 8.30 (0.99) | 100 | 107.21 (15.62) | Non-verbal Fluency (Ideational) | -0.87 |
| | 18 TD | 8.56 (1.00) | 83.3 | 110.83 (10.38) | | |
| Craig and Baron-Cohen (1999) | 30 Autism/Asp | 12.9 (3.1) | NR | NR | Non-verbal Fluency (Ideational) | -1.35 |
| | 15 TD | 5.20 (2.7) | NR | NR | | |
| Van Eylen et al. (2015) | 50 ASD | 12.22 (2.58) | 60 | 104.32 (10.83) | Non-verbal Fluency (Design) | -0.24 |
| | 50 TD | 12.48 (2.72) | 60 | 107.72 (9.3) | Non-verbal Fluency (Ideational) | -0.66 |
| Mostert-Kerckhoffs et al. (2015)[b] | 32 Autism/Asp | 11.30 (1.4) | 77.5 | 110.6 (1.4) | Non-verbal Fluency (Design) | -0.88 [b] |
| | 27 TD | 11.00 (1.2) | 81.2 | 112.5 (14.5) | | |

| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size (d) |
|---|---|---|---|---|---|---|
| Low et al. (2009) | 27 Autism/Asp | 8.26 (2.17) | 85.2 | NR | Non-verbal Fluency (Ideational) | 0.55 |
| | 27 TD | 6.60 (1.31) | 85.2 | NR | | |

The ASD sample is described in the top row and the typically developing control sample in the bottom row. Negative effect sizes indicate that the ASD sample performed less well on generativity tasks than typically developing individuals, in line with predicted outcome.
[a] Calculation based on data provided by the authors
[b] Calculation based on means and standard deviations estimated from graphs presented in articles
NR: Not reported in the article or insufficient information available
Asp: Asperger's syndrome
HFA: High Functioning Autism
ASD: Autistic Spectrum Disorder

**Table 1.4** Studies Comparing Generativity in Samples of Adults with ASD and TD Controls

| **Article** | **Sample** | **Age (years) (M±SD)** | **Gender (% male)** | **IQ (M±SD)** | **Type of Generativity Task** | **Effect Size *(*d)** |
|---|---|---|---|---|---|---|
| Bramham et al. (2009) | 34 Asp/Atypical | 32.76 (12.47) | 80 | 107 (16.38) | Verbal Fluency (Phonemic) | -0.76 |
| | 31 TD | 32.81 (9.02) | 66.5 | 109.84 (16.7) | | |
| Ambery et al. (2006) | 27 Asp | 37.6 (14.6) | 81.5 | 106.1 (15.7) | Verbal Fluency (Phonemic) | -0.61 |
| | 20 TD | 33.50 (12) | 80.0 | 107.05 (13.1) | | |
| Inokuchi and Kamio (2013) | 30 ASD/Asp | 19.2 (2.6) | 83.3 | 99.6 (12.8) | Verbal Fluency (Phonemic) | -0.31 |
| | 18 TD | 20.10 (2) | 83.3 | 101.9 (13.9) | Verbal Fluency (Semantic) | -0.60 |
| Kleinhans et al. (2008) | 14 ASD/Asp | 23.79 (9.58) | 100 | 98.14 (11.84) | Verbal Fluency (Phonemic) | -2.23 |
| | 14 TD | 22.41 (8.67) | 100 | 113.43 (13.91) | Verbal Fluency (Semantic) | -1.07 |
| Lever and Geurts (2016) | 118 ASD | 47.6 (14.9) | 70.3 | 114.8 (16.9) | Verbal Fluency (Phonemic) | -0.32 |
| | 118 TD | 47.70 (15.4) | 70.3 | 114.3 (15.3) | Verbal Fluency (Semantic) | -0.32 |

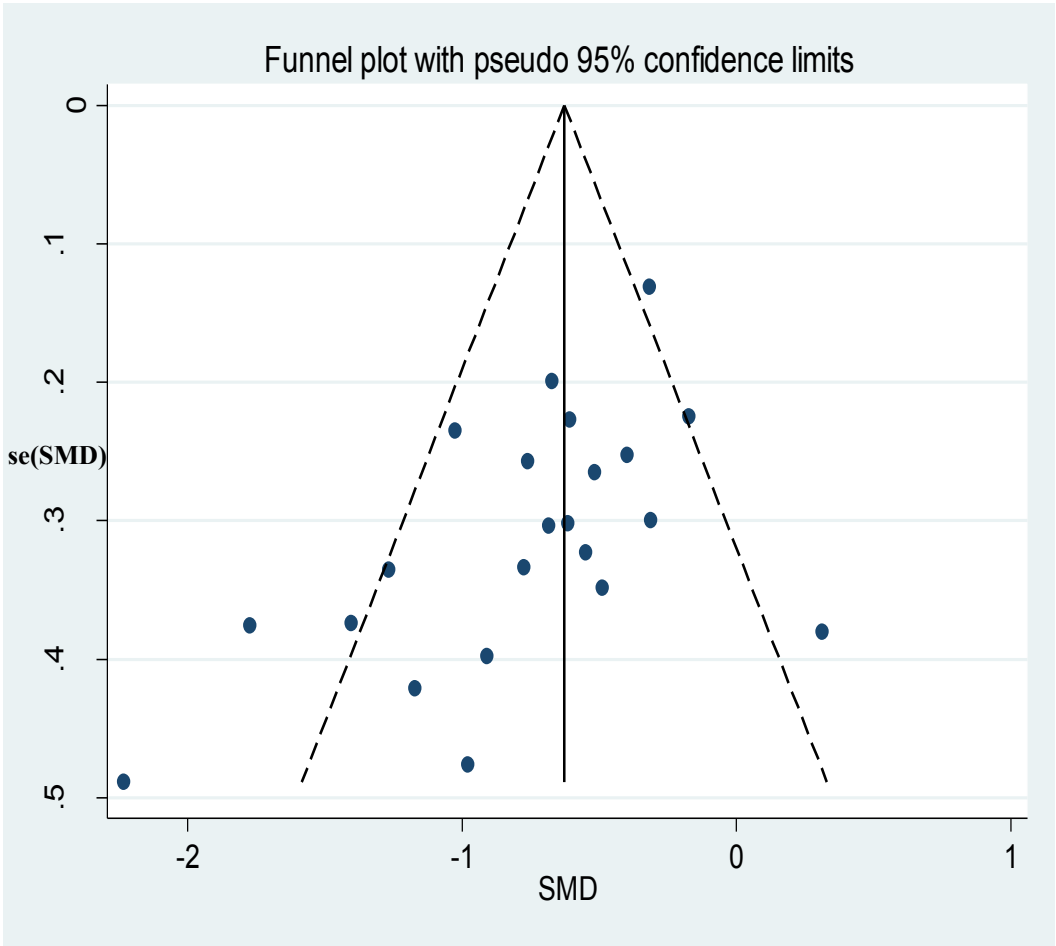
| Article | Sample | Age (years) (M±SD) | Gender (% male) | IQ (M±SD) | Type of Generativity Task | Effect Size *(*d) |
|---|---|---|---|---|---|---|
| Lind and Bowler (2010) | 14 HFA | 41.38 (12.71) | 78.6 | 105.86 (14.52) | Verbal Fluency (Phonemic) | 0.31 |
| | 14 TD | 43.83 (10.39) | 78.6 | 108.57 (18.2) | Verbal Fluency (Semantic) | -0.01 |
| | | | | | Non-verbal Fluency (Ideational) | -0.36 |
| Spek et al. (2009) | 62 HFA/Asp | 39.67 (11.41) | 91.9 | 113.33 (14.57) | Verbal Fluency (Phonemic) | -0.61 |
| | 30 TD | 39.89 (11.45) | 93.3 | 116.77 (11.33) | Verbal Fluency (Semantic) | -0.89 |
| Rumsey and Hamburger (1988) | 10 Autism | 26.4 (7.35) | 100 | 103.4 (9.47) | Verbal Fluency (Phonemic) | -0.98 |
| | 10 TD | 28.40 (4.86) | 100 | 112.8 (3.97) | | |
| Geurts and Vissers (2012) | 23 ASD/Asp | 63.6 (7.5) | 78.3 | 109.2 (10.3) | Verbal Fluency (Phonemic) | -0.68 |
| | 23 TD | 63.70 (8.1) | 78.3 | 109.8 (7.9) | | |
| Lopez et al (2005) | 17 Autism | 29.1 (8.0) | 82.4 | 77.0 (15.0) | Verbal Fluency (Phonemic) | -0.49 |
| | 17 TD | 29.4 (11.4) | 64.7 | 89.0 (13.0) | | |

The ASD sample is described in the top row and the typically developing control sample in the bottom row. Negative effect sizes indicate that the ASD sample performed less well on generativity tasks than typically developing individuals, in line with predicted outcome.
NR: Not reported in the article or insufficient information available
Asp: Asperger's syndrome; HFA: High Functioning Autism; ASD: Autistic Spectrum Disorder

Publication bias was assessed through visual inspection of a funnel plot (figure 1.3) which suggested some asymmetry. This was confirmed statistically using the Eggers test ($t = -2.47$, $p<0.01$); although the large degree of heterogeneity between studies should be taken into account (Ioannidis & Trikalinos, 2007).

Calculation of Rosenthal's Failsafe-*N* for both the child and adult studies ranged from 148 to 693. It is therefore possible that the current analysis overstates the poorer performance of individuals with ASD on verbal fluency tasks. However, as a large number of studies would be required to render the findings insignificant, it is highly unlikely that the observed effect is an artefact of publication bias.

**Figure 1.3** Funnel plot for child and adult studies comparing phonemic fluency tasks between ASD and TD.



### *5.2.2 Semantic Fluency Tasks*

The results of the combined and subgroup analysis of both the child and adult studies for category fluency tasks are presented in Figure 1.4. The overall effect of ASD on verbal generativity as assessed by category fluency tasks was estimated to be -0.63 (95% CI = -0.83 - -0.43) demonstrating a medium sized effect (Cohen, 1992). This was highly significant ($p<0.001$).

Meta-analysis of the twenty child studies ($N_{ASD}$ and $N_{TD}$ = 567) looking at generativity performance as measured by semantic fluency tasks, estimated the population effect size as -0.65 (CI = -0.89 - -0.41, $p<0.001$). This was significant and suggests that children with ASD show a medium impairment on verbal generativity performance in tasks using semantic cues (Cohen, 1988). For the five adult samples ($N_{ASD}$ = 238 and $N_{TD}$ = 194) there was an estimated medium effect size of -0.55 (CI = -0.89 – -0.22) in the predicted direction indicating that adults with ASD typically performed worse on measures of semantic generativity ($p=0.001$).

There was a large degree of variability in effect size across the studies ranging from no effect to large effects. Analysis of heterogeneity was substantial for the child studies ($Q = 70.21$, $df = 19$, $p<0.01$, $I = 72.9\%$) but not the adult studies ($Q = 8.54$, $df = 4$, $p = 0.07$, $I = 53.2\%$). Analysis of publication bias using a funnel plot indicated only a small degree of spread (Figure 1.5). Further analysis did not provide evidence for a publication bias (Eggers test, $t = -1.69$, $p = 0.10$).

**Figure 1.4** Forest plot for studies comparing generativity performance in semantic fluency tasks between ASD and TD.

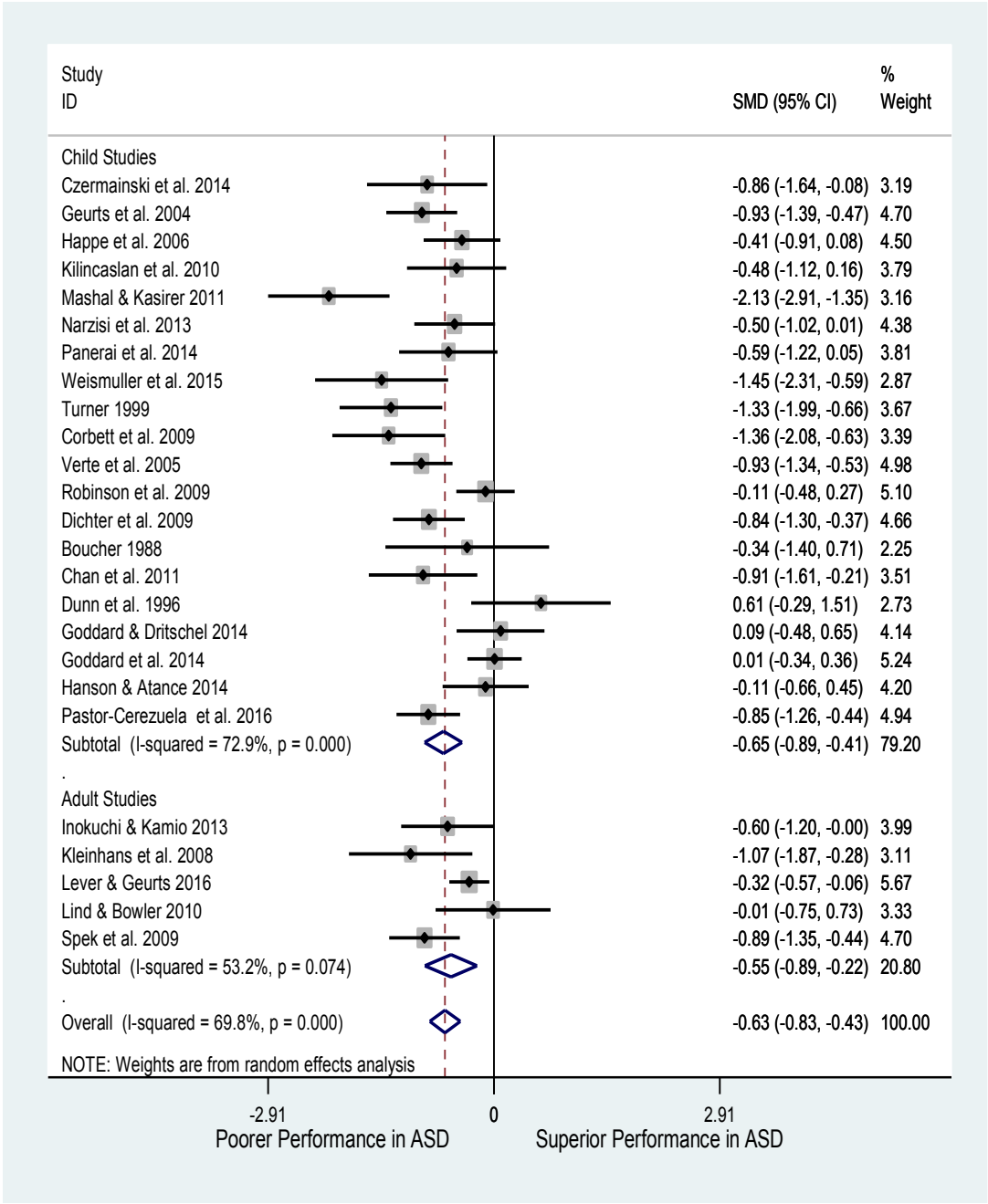
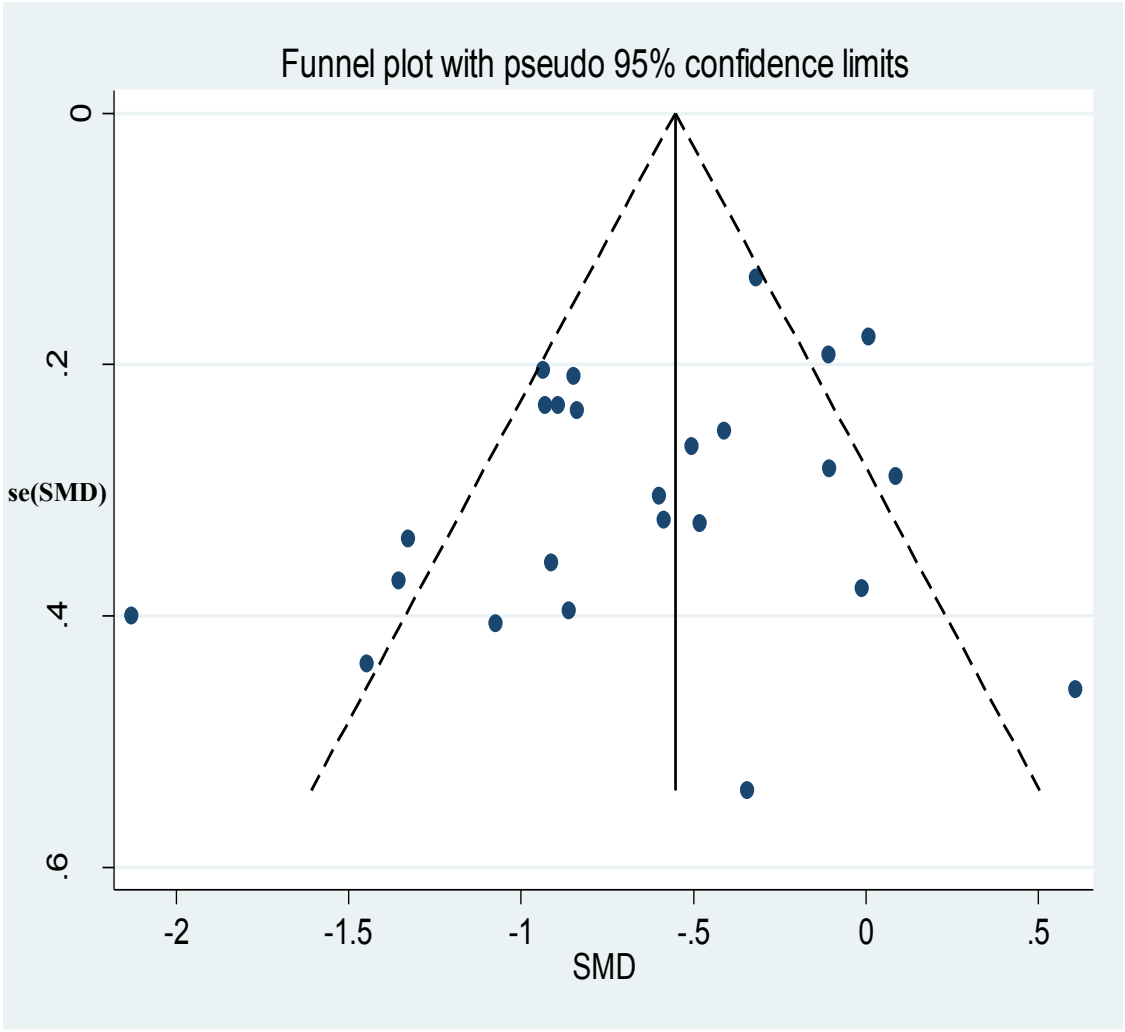


**Figure 1.5.** Funnel plot for child and adult studies comparing semantic fluency tasks between ASD and TD.



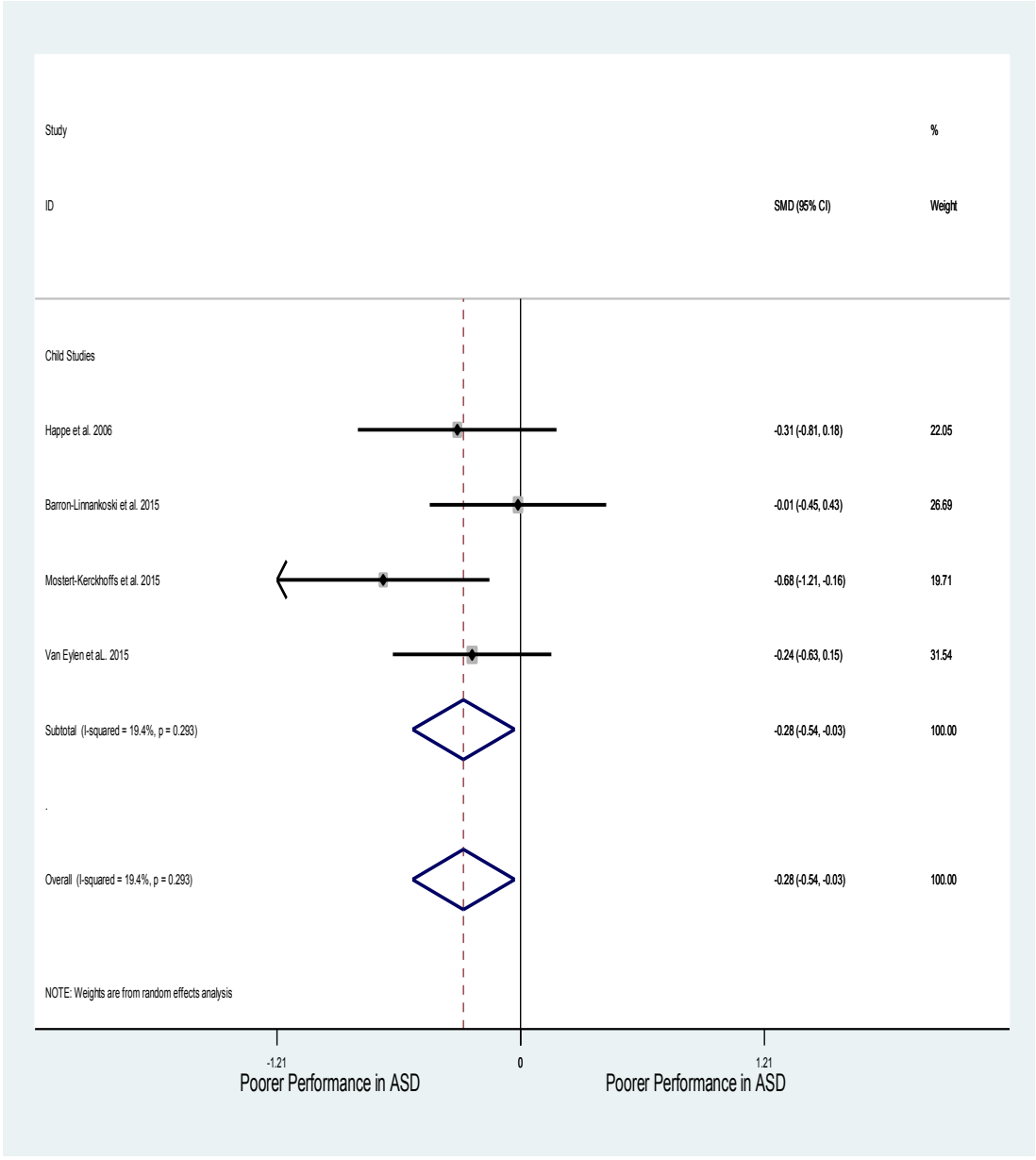

## 5.3 Non-verbal Generativity

### *5.3.1 Design Fluency*

Four studies reported performance on design fluency tasks, all of which used child and adolescent participants. The samples included 144 children with ASD (mean age = 10.88, SD = 1.31) and 169 TD controls (mean age = 10.95, SD = 1.39). As can be seen in Figure 1.6, the overall effect size was estimated to be -0.28 (CI = -0.54 to -

0.03, p=0.029) indicating a small, marginally significant effect. The degree of heterogeneity across studies was insignificant (Q = 3.72, df = 3, p = 0.293, $I^2$ = 19.4).

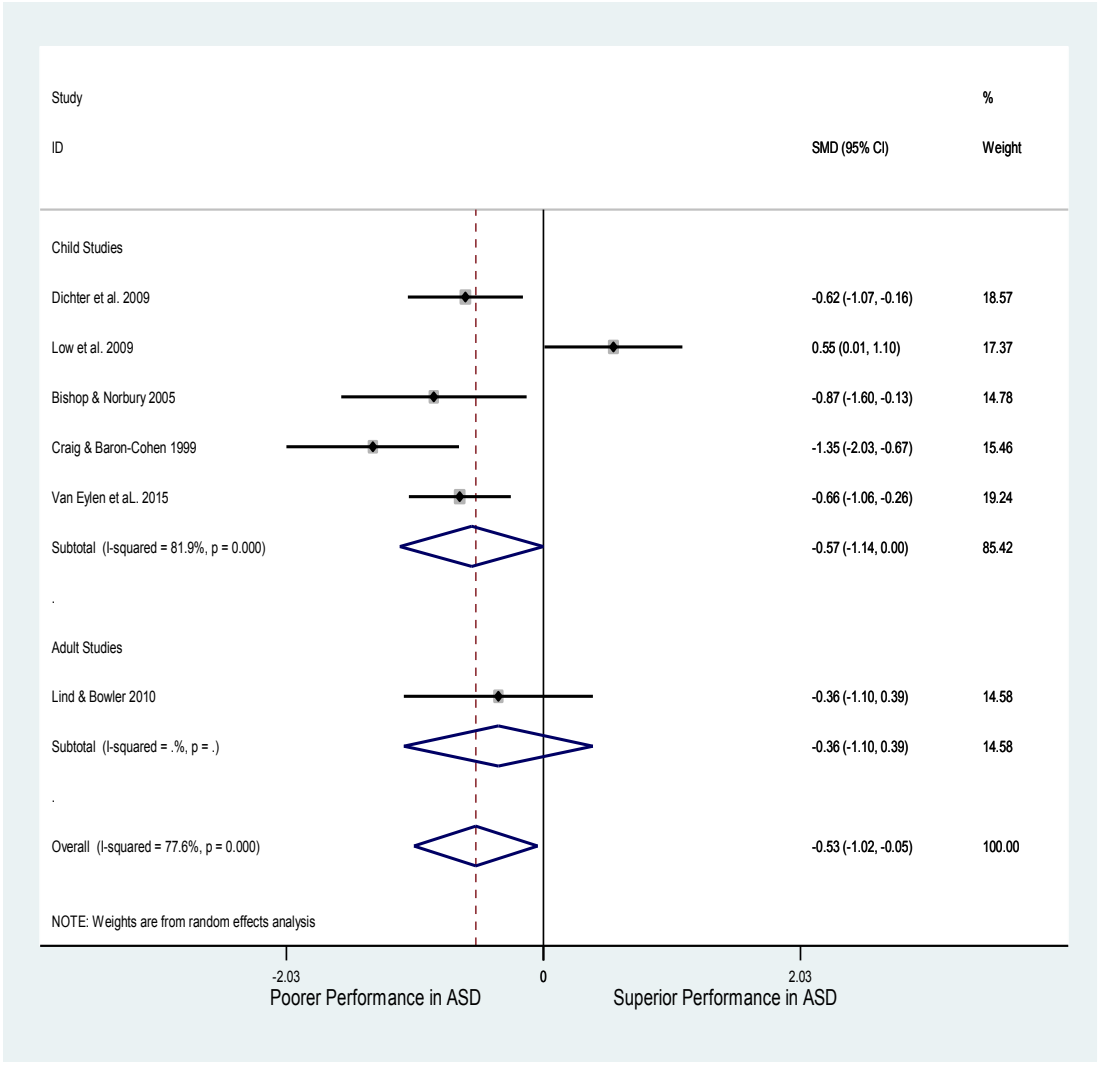
**Figure 1.6** Forest plot for studies comparing generativity performance in design fluency tasks between ASD and TD.



### *5.3.2 Ideational Fluency*

There were five child studies and one adult study reporting generativity performance as measured through ideational fluency tasks (ES = -0.53, CI = -1.02 – -0.05, p=0.032). Sub-group analysis estimated the population effect size amongst children as -0.57 (CI = -1.14 – 0.00, p=0.05). Heterogeneity was large (Q = 22.28, df = 5 p<0.01, $I^2$ = 77.6%) (Figure 1.7).

**Figure 1.7** Forest plot for studies comparing generativity performance in ideational fluency tasks between ASD and TD



Tests of asymmetry to assess publication bias were not applied to the design and ideational fluency analyses as the power for these was too low to provide a meaningful analysis (Higgins & Green, 2008).

### *5.3.3 Analysis of Potential Effect Modifiers*

Due to significant heterogeneity, further sub-group analyses were conducted for phonemic and semantic fluency in the child samples; and phonemic fluency in the adult samples. Ideational generativity was not analysed further due to the small number of studies. Categorical analysis was performed as there were less than or equal to 20 studies in each group, which precludes meta-regression. This method has been followed elsewhere (Wykes et al., 2011). Table 1.5 displays the results of the analyses.

### *5.3.4 IQ*

Studies that reported IQ of their clinical samples were grouped based on whether the ASD participants had a 'Lower' or 'Higher' IQ relative to the overall mean. Sub-group comparisons indicated that effect sizes were greater in those studies whose ASD participants had a lower IQ. This was true for both types of task and age group. The $I^2$ statistic suggested that heterogeneity remained moderate to substantial for the lower IQ group in the adult phonemic studies and higher IQ group in the child semantic studies only.

### *5.3.5 Gender Ratio*

For each task comparison, studies were grouped based on whether they had a higher or lower proportion of males relative to the overall mean. There was a high proportion of male participants across all the studies included in the meta-analyses (mean 88.80%, SD = 10.34). Findings were mixed, with child semantic and adult phonemic fluency tasks showing greater effect size when more female participants

were included in the sample. However, the opposite was true for phonemic fluency in the child studies. Likewise, heterogeneity varied across these groups.

### *5.3.6 Study Quality*

To explore the impact of study quality on effect size, studies were grouped based on their overall quality score (mean = 18.35, SD = 2.54) and their methodological score (mean = 10.84, SD = 2.23) as per the study quality assessment tool. Those that scored below the mean (overall score < 19 or methodological score < 11) were rated as 'lower quality' and above the mean as 'higher quality'. Study ratings and effect size estimates were consistent for both measures and so only the group comparisons based on overall quality are shown. In all cases, studies that were rated as lower quality shared an effect size that was greater than those studies rated as higher quality. For phonemic tasks the only group to show significant heterogeneity was the lower quality adult studies. However, both groups in the child semantic fluency studies had an $I^2$ statistic indicating substantial heterogeneity

Table 1.5 *Moderator Analyses for Phonemic Fluency (child and adult studies) and Semantic Fluency (child studies)*

| Moderator Analysis | | N. Studies | SMD | 95% confidence intervals | Heterogeneity $\chi^2$ | p | $I^2$ |
|---|---|---|---|---|---|---|---|
| ***Phonemic – Child Studies*** | | | | | | | |
| Study Quality | Lower Quality | 7 | -1.045 | -1.405 - -0.686 | 11.54 | 0.073 | 48.0% |
| | Higher Quality | 5 | -0.597 | -0.897 - -0.297 | 7.91 | 0.095 | 49.4% |
| Composite IQ | Lower IQ | 7 | -0.824 | -1.086 - -0.561 | 9.64 | 0.141 | 37.7% |
| | Higher IQ | 3 | -0.629 | -1.219 - -0.038 | 5.33 | 0.070 | 62.5% |
| Gender | Lower Males | 6 | -0.870 | -1.347 - -0.392 | 16.93 | 0.005 | 70.5% |
| | Higher Males | 6 | -0.791 | -1.073 - -0.509 | 8.28 | 0.141 | 39.6% |
| ***Phonemic –Adult Studies*** | | | | | | | |
| Study Quality | Lower Quality | 3 | -1.224 | -2.169 - -0.280 | 7.97 | 0.019 | 74.9% |
| | Higher Quality | 7 | -0.433 | -0.644 - -0.223 | 7.70 | 0.261 | 22.1% |
| Composite IQ | Lower IQ | 4 | -0.938 | -1.712 - -0.164 | 12.14 | 0.007 | 75.3% |
| | Higher IQ | 6 | -0.469 | -0.716 - -0.223 | 7.88 | 0.163 | 36.6% |
| Gender | Lower Males | 7 | -0.422 | -0.636 - -0.208 | 7.41 | 0.285 | 19.0% |
| | Higher Males | 3 | -1.208 | -2.154 - -0.262 | 9.13 | 0.010 | 78.1% |

| Moderator Analysis | | N. Studies | SMD | 95% confidence intervals | $\chi^2$ | p | $I^2$ |
|---|---|---|---|---|---|---|---|
| ***Semantic – Child Studies*** | | | | | | | |
| Study Quality | Lower Quality | 11 | -0.819 | -1.227 - -0.410 | 35.83 | <.001 | 72.1% |
| | Higher Quality | 9 | -0.498 | -0.783 - -0.213 | 29.22 | <.001 | 72.6% |
| Composite IQ | Lower IQ | 12 | -0.804 | -1.166 - -0.442 | 8.70 | 0.069 | 54.0% |
| | Higher IQ | 5 | -0.503 | -0.798 - -0.209 | 39.92 | <.001 | 72.4% |
| Gender | Lower Males | 9 | -0.429 | -0.735 - -0.122 | 24.98 | 0.002 | 68.0% |
| | Higher Males | 11 | -0.846 | -1.166 - -0.526 | 29.96 | 0.001 | 66.6% |

# 6 Discussion

The question of whether generativity is impaired in individuals with ASD is one that has not yet been fully answered within the literature. This paper sought to determine whether an overall impairment is evident in children and adults with ASD and articulate the magnitude of effect through quantitative synthesis. To be as comprehensive as possible, it included studies that utilised a range of verbal and non-verbal measures of mental fluency namely; phonemic verbal fluency, semantic verbal fluency, ideational fluency and design fluency. The fact that individual study effects overlapped the composite effect sizes indicated that combining measures of generativity was warranted and did not enhance the overall effect.

## 6.1 Generativity Impairment in Autism

The findings of the meta-analyses provide strong support for the notion that individual's with ASD perform significantly worse on tasks of generativity than TD controls. There was also evidence that these impairments extend to both verbal and non-verbal tasks of generativity. These findings are in line with the conclusions of previous narrative reviews (Hill, 2004a, 2004b) and may lend support to the executive dysfunction hypothesis of ASD (Russell, 1997).

A medium sized effect was found for both phonemic generativity (generating words based on letter cues) and semantic generativity (words based on category exemplars such as animals, fruits etc.). This finding is interesting given that phonemic and semantic fluency tasks have been linked to different areas of the brain which may rely on different mechanisms (Henry & Bettenay, 2010). Further, there has been the suggestion that phonemic tasks are more sensitive to executive dysfunction as they place greater demands on cognitive processes than semantic ones (Perret, 1974; Shao

et al., 2014). Although the current review did indeed find a slightly larger effect for tasks using letter based cues, there was also suggestion of impairment on semantic tasks. This finding is one that has been replicated in individuals with frontal focal cortical lesions, suggesting a similar frontal executive impairment in ASD (Henry & Crawford, 2004).

There have been fewer studies using non-verbal measures of generativity. However, the small number of studies that did include design or ideational fluency tasks (n=10) demonstrated a small (design) to medium (ideational) effect, with TD controls outperforming those with ASD. Previous reviews have linked this to both an impairment in generating novel responses and behaviours, and a failure of self-regulation through inhibition and self-monitoring (Hill, 2004a; Turner, 1999).

It has been argued that previous contradictory findings within the generativity literature are because test requirements vary with regards to their cognitive load (Inokuchi & Kamio, 2013). For example, those tasks that provide category cues are thought to be less cognitively demanding than those that require participants to generate a truly de novo response. Yet results of the current review suggested a trend towards verbal tasks showing larger effects than non-verbal ones. This is surprising given that non-verbal generativity tasks tend to be more "open-ended" in nature; something which has been linked with greater impairment in those with ASD (White, Burgess, & Hill, 2009).

## 6.2 Generativity Performance and Age

The similarity in effect size estimates for both child and adult populations is again an interesting one. It suggests that, in general, performance remains relatively

stable across time which is a finding that has been replicated elsewhere (Lever & Geurts, 2016). This is despite studies of non-ASD populations finding that generativity, in particular verbal fluency, declines with age (Brickman et al., 2005; Clark et al., 2009). This may be a consequence of the 'safeguard hypothesis', whereby ASD moderates the typical age-related patterns of cognitive decline seen in those without the disorder (Geurts, 2016). In fact, our review found a small difference in effect sizes between the child and adult studies suggesting that generativity impairments may be more evident in childhood. However, the studies lacked sufficient power to draw any firm conclusions on this.

### 6.3 Heterogeneity and Potential Effect Modifiers

Given that there was significant heterogeneity amongst the studies, findings may overestimate the actual effect of generativity impairments in ASD (Higgins & Green, 2008). Subgroup comparisons indicated a potential moderating effect of study quality; with 'lower quality' studies producing larger effect sizes. These papers had smaller numbers of participants; lacked power; failed to match samples; and did not include diagnostic screens to confirm clinical diagnoses. Although these findings should be interpreted cautiously (Higgins & Green, 2008), they suggest that generativity deficits may be inflated by poor study design. Indeed, those papers that were rated as poorest quality according to the quality assessment tool (Kleinhans, Muller, Cohen, & Courchesne, 2008; Turner, 1999b) demonstrated some of the largest effect sizes. Arguably, more methodologically sound studies of generativity are needed to draw concrete conclusions about a generativity deficit in ASD.

Contrasts based on IQ indicated that effect sizes may differ for children with ASD depending on whether they have a lower or higher IQ relative to the mean. Studies that included ASD children with a lower IQ tended to yield larger effect sizes. This is not surprising given that IQ is a strong predictor of fluency performance (Arffa, 2007; Pastor-Cerezuela et al., 2016), so it would be expected that those children with a poorer IQ perform less well on measures of generativity. However, it raises the questions of whether impaired performance in ASD is a function of lower IQ rather than of a generativity deficit per se. This is further evidence as to a need for greater methodological rigour.

Subgroup analyses of gender did not present any notable considerations. However, from a more observational perspective, many of the studies across the generativity literature lack sufficient female participants. All the studies had between 70-100% males in their samples. This may reflect the increased prevalence of ASD in males (Baird, 2006). However, it also calls into question the generalizability of findings given that it has been consistently shown that males and females have differing executive function abilities (Capitani et al., 1999; Kimura, 1992). One question that remains unanswered is whether gender differences in generativity skills may explain the observed differences in autism symptomology, such as reduced repetitive behaviours in females (Mandy et al., 2012).

### 6.4 Limitations of the Review

In addition to accounting for heterogeneity through moderator variable analysis, it is also important to consider theoretical reasons why such variability might exist. It is questionable whether generativity is a discrete enough entity to be accurately

measured experimentally. Some argue that it is a specific neurocognitive process that can be captured through fluency tasks (Dichter et al., 2009). However, executive control is inherently complex and it is therefore difficult to isolate it's subcomponents (Kenworthy et al., 2008). Several studies suggest that additional processes are called upon when completing fluency tasks. These include; response selection and inhibition, focused attention, verbal short-term memory and sustained attention (Goddard, Dritschel, & Howlin, 2014; Happé, Booth, Charlton, & Hughes, 2006; Pastor-Cerezuela et al., 2016). Fluency performance is also affected by processing speed and motor performance (Spek et al., 2009) which is significant given that motor difficulties are more common in children with ASD (Ming, Brimacombe, & Wagner, 2007). Therefore, it is possible that fluency tasks measure a range of neurocognitive processes and, as such, poor performance is not necessarily indicative of a generativity deficit.

Similarly fluency tasks have been described as a hybrid of measures that lack construct validity (Shao et al., 2014). For example, Boucher (1988) found that those with ASD were unimpaired on verbal fluency tasks using specific cues but performed worse than controls when the tasks became more open-ended as would be expected in "real-life". Further, there are subtle differences in the ways in which the tasks have been administered across studies. A clear example of this is in the study by Craig & Baron-Cohen (1999) (included in the current review) who used a hybrid of the Use of Objects task and asked children to generate ideas as to what a foam shape might be. Interestingly, this study gave rise to large effect sizes compared with other studies of ideational fluency, which calls in to question how comparable different measures of fluency are.

The current review included studies where participants had an IQ within the normal range to limit the number of confounding variables. Population-based studies have indicated that there is an increased prevalence of intellectual disability in those with Autism (Deb & Prasad, 1994). Therefore, the current review may only be generalizable to an ASD population with a normal range of intellectual functioning. It was also beyond the scope of the current review to consider how co-morbid disorders such as ADHD impact on generativity performance. This might be an important future direction given the prevalence of co-morbidity within this population (Kohane et al., 2012).

Given that generativity is a neglected aspect of executive function within the literature, the current review sought to be as comprehensive as possible. In doing so it included results of more than one generativity task for some studies (e.g. semantic and phonemic fluency) and as such violated the assumption of statistical independence. However, this is not uncommon amongst the meta-analytic literature and as argued elsewhere, there was no elegant way around this problem (Henry & Crawford, 2004). Despite this, there were still a limited number of studies that explored generativity in children and adults with autism, particularly when compared to other areas of executive function such as visuo-spatial performance (e.g. Muth et al., 2014). This limited the power of the review.

There are many threats to the validity of a meta-analysis. We attempted to address these by adhering to strict inclusion and exclusion criteria in addition to only including peer reviewed publications (Sharpe, 1997). This meant that a number of key papers reporting non-significant findings were excluded due to their use of a mixed sample based on age (Hill & Bird, 2006; Minshew, Goldstein, Muenz, & Payton, 1992;

Minshew, Goldstein, & Siegel, 1995). There was also evidence of a publication bias in the case of the phonemic fluency tasks. By not including unpublished works it was possible that the current review failed to account for the 'file-drawer effect' (Rosenthal, 1979) but conversely, including unpublished studies of poorer quality would not have been sufficient to address all validity concerns (Sharpe, 1997). Regardless, these factors should be considered when interpreting the findings.

### 6.5 Clinical Implications and Suggestions for Future Research

In conclusion, our review found overall support for impaired performance on generativity tasks in children and adults with ASD. Generativity abilities have been implicated in memory, pretend play, communication and other complex neurocognitive processes (Bishop & Norbury, 2005; Dichter et al., 2013; Goddard, Dritschel, Robinson, et al., 2014; Rutherford & Rogers, 2003). Therefore, impairments may be considered in the context of an executive dysfunction hypothesis of autism. The current review does not explore in detail the specific mechanisms that might be at play when using fluency tasks. The ways in which individuals with ASD cluster and switch their responses and the prototypicality of their responses, may be more informative in our understanding of ASD profiles than looking at overall generation of responses alone (Beacher et al., 2012; Dunn, Gomes, & Sebastian, 1996). Indeed generativity tasks, like other measures of executive function, may fail to capture the complexity of the generativity impairments that those with ASD experience in the "real-world" (Kenworthy et al., 2008; Leung & Zakzanis, 2014a; White et al., 2009). This along with sample characteristics and poor study design may account for the heterogeneity seen amongst the current literature.

# 7 References

American Psychiatric Association. (1987). *DSM-III. Diagnostic and Statistical Manual of Mental Disorders, 3rd edition* (3rd, revis ed.). Washington, DC, US: Author.

American Psychiatric Association. (2000). *DSM-IV*. *Diagnostic and Statistical Manual of Mental Disorders 4th edition TR*. Washington, DC,: Author.

American Psychiatric Association. (2013). *DSM-V*. *Diagnostic and Statistical Manual of Mental Disorders 5th Edition*. Washington, DC,: Author.

Anderson, V. (1998). Assessing Executive Functions in Children: Biological, Psychological, and Developmental Considerations. *Neuropsychological Rehabilitation*, *8*(3), 319–349.

Ardila, a, Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *15*(1), 31–6.

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, *22*(8), 969–978.

Baird, G. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215.

Barron-Linnankoski, S., Reinvall, O., Lahervuori, A., Voutilainen, A., Lahti-Nuuttila, P., & Korkman, M. (2015). Neurocognitive performance of children with higher functioning autism spectrum disorders on the NEPSY-II. *Child Neuropsychology*,

*21*(1), 55–77.

Beacher, F. D., Radulescu, E., Minati, L., Baron-Cohen, S., Lombardo, M. V., Lai, M. C., … Critchley, H. D. (2012). Sex differences and Autism: Brain function during verbal fluency and mental rotation. *PLoS ONE*, *7*(6), 1–16.

Begeer, S., Wierda, M., Scheeren, A. M., Teunisse, J.-P., Koot, H. M., & Geurts, H. M. (2014). Verbal fluency in children with autism spectrum disorders: Clustering and switching strategies. *Autism*, *18*(8), 1014–1018.

Benton, A., & Hamsher, K. (1983). Multilingual Aphasia Examination. PAR Inc.

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. *Neuropsychologia*, *6*(137), 53–60. http://doi.org/10.1016/0028-3932(68)90038-9

Bereiter, C., Harris, C. W., Archer, E. J., & Klausmeier, H. J. (1960). Verbal and Ideational Fluency in Superior Tenth Grade Students. *Journal of Educational Psychology*, *61*(6), 337–345.

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. 2: Response inhibition. *Autism : The International Journal of Research and Practice*, *9*(1), 29–43.

Bordignon, S., Endres, R. G., Trentini, C. M., & Bosa, C. A. (2015). Memory in children and adolescents with autism spectrum disorder: A systematic literature review. *Psychology & Neuroscience*, *8*(2), 211–245.

Boucher, J. (1988). Word fluency in high-functioning autistic children. *Journal of Autism and Developmental Disorders*. University of Warwick, Coventry United Kingdom: Plenum Publishing Corp.

Brickman, A. M., Paul, R. H., Cohen, R. A., Williams, L. M., MacGregor, K. L.,

Jefferson, A. L., … Gordon, E. (2005). NIH Public Access. *Archives of Clinical Neuropsychology*, *20*(5), 561–573.

Bristow, E. (2016). *Developing and Piloting a New Measure of Executive Functioning for Children with Autism Spectrum Disorder (ASD)*. University College London.

Capitani, E., Laiacona, M., Barbarotto, R., Hospital, S. P., Unit, N., & Foundation, S. M. (1999). Gender affects word retrieval of certain category in semantic fluency tasks. *Cortex*, *35*(1998), 273–278.

Chan, D. W., Cheung, P.-C., Lau, S., Wu, W. Y., Kwong, J. M., & Li, W.-L. (2001). Assessing ideational fluency in primary students in Hong Kong. *Creativity Research Journal*, *13*(3–4), 359–365.

Clark, L. J., Gatz, M., Zheng, L., Chen, Y.-L., McCleary, C., & Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, *24*(6), 461–8.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences*. http://doi.org/10.1234/12345678

Cohen, J. (1992). Cohen 1992 Power Primer.pdf. *Psychological Bulletin*, *112*(1), 155–159.

Craig, J., & Baron-Cohen, S. (1999). Creativity and imagination in autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *29*(4), 319–326.

Critical Appraisal Skills Programme (CASP). (2014). Case Control Study Checklist. Retrieved from http://www.casp-uk.net

Crowe, S. F. (1998). Decrease in Performance on the Verbal Fluency Test as a

Function of Time: Evaluation in a Young Healthy Sample. *Journal of Clinical and Experimental Neuropsychology*, *3395*(3), 391–401.

Deb, S., & Prasad, K. B. (1994). The prevalence of autistic disorder among children with a learning disability. *The British Journal of Psychiatry*, *165*(3), 395–399.

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2008). Statistical Methods for Examining Heterogeneity and Combining Results from Several Studies in Meta-Analysis. In *Systematic Reviews in Health Care: Meta-Analysis in Context: Second Edition* (pp. 285–312).

Delis, D., Kaplan, E., & Kramer, J. (2001). Delis-Kaplan executive function system (D-KEFS). *Canadian Journal of School Psychology*, *20*(1–2), 117–128.

Dichter, G. S., Drive, M., Hill, C., Lam, K. S. L., Drive, M., Hill, C., … Bodfish, J. W. (2013). NIH Public Access, *39*(9), 1298–1304.

Dichter, G. S., Lam, K. S. L., Turner-Brown, L. M., Holtzclaw, T. N., & Bodfish, J. W. (2009). Generativity abilities predict communication deficits but not repetitive behaviors in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(9), 1298–1304.

Dunn, M., Gomes, H., & Sebastian, M. (1996). Prototypicality of responses of autistic, language disordered, and normal children in a word fluency task. *Child Neuropsychology*. United Kingdom: Taylor & Francis.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634.

Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge: Cambridge

University Press.

Field, A. P. (2000). A Bluffer's Guide to Meta-Analysis. *Newsletter of the Mathematical, Statistical and Computing Section of the British Psychological Society*, *7*, 16–25.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *The British Journal of Mathematical and Statistical Psychology*, *63*(Pt 3), 665–694.

Fombonne, E., Quirke, S., & Hagen, A. (2011). Epidemiology of pervasive developmental disorders. In D. . Amaral, G. Dawson, & D. H. Geschwind (Eds.), *Autism Spectrum Disorders* (pp. 90–111). New York: Oxford University Press.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. J. (1991). A PET study of word finding. *Neuropsychologia*, *29*(12), 1137–1148.

Gaillard, W. D., Hertz-Pannier, L., Mott, S. H., Barnett, a S., LeBihan, D., & Theodore, W. H. (2000). Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology*, *54*(1), 180–185.

Geurts, H. M. (2016). Autism and Age Related Cognitive Decline. In S. D. Wright (Ed.), *Autism Spectrum Disorder in Mid and Later Life* (pp. 154–162). London: Jessica Kingsley Publishers.

Goddard, L., Dritschel, B., & Howlin, P. (2014). A preliminary study of gender differences in autobiographical memory in children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(9), 2087–2095.

Goddard, L., Dritschel, B., Robinson, S., & Howlin, P. (2014). Development of autobiographical memory in children with autism spectrum disorders: Deficits, gains, and predictors of performance. *Development and Psychopathology*, *26*(1), 215–228.

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.

Happe, F., Fleminger, S., & David, A. S. (2006). Intact theory of mind in TBI with behavioural disturbance. *Brain and Cognition*, *60*(2), 196–198.

Harris, R. J., Deeks, J. J., Altman, D. G., Bradburn, M. J., Harbord, R. M., & Sterne, J. A. C. (2008). metan: fixed- and random-effects meta-analysis. *The Stata Journal*, *8*(1), 3–28.

Hedges, L. V. (1992). Meta-analysis. *JOurnal of Educational Statistics*, *17*(4), 279–296.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. *Phytochemistry*, *72*(13), 369. http://doi.org/10.1016/j.phytochem.2011.03.026

Henry, J., & Crawford, J. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*(2), 284–295.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*(2), 284–295.

Henry, L. a., & Bettenay, C. (2010). The Assessment of Executive Functioning in Children. *Child and Adolescent Mental Health*, *15*(2), 110–119.

Higgins, J., & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of. The Cochrane Collaboration* (Vol. Version 5.).

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558.

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), 2822–35.

Hocevar, D. (1979). Ideational fluency as a confounding factor in the measurement of originality. *Journal of Educational Psychology*, *71*(2), 191–196.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, *11*(2), 193–206.

Inokuchi, E., & Kamio, Y. (2013). Qualitative analyses of verbal fluency in adolescents and young adults with high-functioning autism spectrum disorder. *Research in Autism Spectrum Disorders*, *7*(11), 1403–1410.

Ioannidis, J. P. A., & Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *CMAJ*, *176*(8), 1091–1096.

Jarrold, C., Boucher, J., & Smith, P. K. (1996). Generativity deficits in pretend play in autism. *British Journal of Developmental Psychology*, *14*(3), 275–300.

Jones-Gotman, M., & Milner, B. (1977). Design fluency: The invention of nonsense drawings after focal cortical lesions. *Neuropsychologia*, *15*(4–5), 653–674.

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), 320–38.

Kimura, D. (1992). Sex differences in the brain. *Scientific American*.

Kleinhans, N. M., Muller, R.-A., Cohen, D. N., & Courchesne, E. (2008). Atypical functional lateralization of language in autism spectrum disorders. *Brain*

*Research*. Elsevier (P.O. Box 211, Amsterdam 1000 AE, Netherlands).

Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Varity of Fields. Retrieved October 26, 2016, from h

Kohane, I. S., McMurry, A., Weber, G., MacFadden, D., Rappaport, L., Kunkel, L., … Churchill, S. (2012). The Co-Morbidity Burden of Children and Young Adults with Autism Spectrum Disorders. *PLoS ONE*, *7*(4), e33224.

Koolen, S., Vissers, C. T. W. M., Egger, J. I. M., & Verhoeven, L. (2014). How stimulus and task complexity affect monitoring in high-functioning adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(10), 2499–2513.

Korkman, M., Kirk, U., & Kemp, S. (1998). NEPSY: A developmental neuropsychological assessment. Psychological Corporation.

Lai, M. C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(1), 11–24.

Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within Autism Spectrum Disorders: What have We Learned from Neuroimaging Studies? *Frontiers in Human Neuroscience*, *7*(October), 733.

Leung, R. C., & Zakzanis, K. K. (2014a). Brief Report : Cognitive Flexibility in Autism Spectrum Disorders : A Quantitative Review. *Journal of Autism and Developmental Disorders*, *44*(10), 2628–2645. http://doi.org/10.1007/s10803-014-2136-4

Leung, R. C., & Zakzanis, K. K. (2014b). Brief report: Cognitive flexibility in autism spectrum disorders: A quantitative review. *Journal of Autism and Developmental Disorders*, *44*(10), 2628–2645.

Lever, A. G., & Geurts, H. M. (2016). Age-related differences in cognition across the adult lifespan in autism spectrum disorder. *Autism Research*, *9*(6), 666–676.

Lezak, M. (1995). *Neuropsychological Assessment*. *Neuropsychological assessment (3rd ed.)*. Oxford: Oxford University Press.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.)*. *Neuropsychological assessment (5th ed.)*. Oxford: Oxford University Press.

Lind, S. E., & Bowler, D. M. (2010). Episodic memory and episodic future thinking in adults with autism. *Journal of Abnormal Psychology*, *119*(4), 896–905.

Low, J., Goddard, E., & Melser, J. (2009). Generativity and imagination in autism spectrum disorder: evidence from individual differences in children's impossible entity drawings. *The British Journal of Developmental Psychology*, *27*(Pt 2), 425–444.

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders*, *42*(7), 1304–1313.

Ming, X., Brimacombe, M., & Wagner, G. C. (2007). Prevalence of motor impairment in autism spectrum disorders. *Brain and Development*, *29*(9), 565–570.

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 Exp Neuropsychol*, *14*(5), 749–761.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology*, *9*(2), 255–261.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & GROUP, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7).

Mostert-Kerckhoffs, M. A. L., Staal, W. G., Houben, R. H., & Jonge, M. V. (2015). Stop and change: Inhibition and flexibility skills are related to repetitive behavior in children and young adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *45*(10), 3148–3158.

Muth, A., Hönekopp, J., & Falter, C. M. (2014). Visuo-Spatial Performance in Autism: A Meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(12), 3245–3263.

Narzisi, A., Muratori, F., Calderoni, S., Fabbro, F., & Urgesi, C. (2013). Neuropsychological profile in high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(8), 1895–1909.

Ozonoff, S., Pennington, Bruce, F., & Rogers, Sally, J. (1991). Executive function deficits in high-functioning autistic.pdf. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *32*(7), 1081–1105.

Pastor-Cerezuela, G., Fernandez-Andres, M.-I., Feo-Alvarez, M., & Gonzalez-Sala, F. (2016). Semantic Verbal Fluency in Children with and without Autism Spectrum Disorder: Relationship with Chronological Age and IQ. *Frontiers in Psychology*, *7*, 921.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*, *12*, 323–330.

Petticrew, M., & Gilbody, S. (2004). Planning and Conducting Systematic Reviews. In S. Michie & C. Abraham (Eds.), *Health Psychology in Practice* (pp. 150–179). Oxford: Blackwell.

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*(3), 362–368.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641.

Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist*, *45*, 775–777.

Rosenthal, R., & Rubin, D. B. (1979). Comparing significance levels of independent studies. *Psychological Bulletin*, *86*(5), 1165–1168. http://doi.org/10.1037//0033-2909.86.5.1165

Russell, J. (1997). *Autism as an executive disorder*. *New York NY US Oxford University Press* (Vol. 313).

Rutherford, M. D., & Rogers, S. J. (2003). Cognitive underpinnings of pretend play in autism. *Journal of Autism and Developmental Disorders*, *33*(3), 289–302.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*(JUL), 1–10.

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.

Snyder, A., Mitchell, J., Bossomaier, T., & Pallier, G. (2004). Creativity Qoutient.pdf. *Creativity Research Journal*, *16*(4), 415–420.

Spek, A., Schatorje, T., & Scholte, E. (2009). Verbal fluency in adults with high functioning autism or Asperger syndrome. *Neuropsychologia*, *47*(3), 652–656.

StataCorp LP. (2011). Stata Statistical Software: Release 12. *2011*.

Torrance, E. P. (1974). Torrance Tests of Creative Thinking. Scholastic Testing Service, Inc.

Turner, M. A. (1999a). Generating novel ideas: Fluency performance in high-functioning and learning disabled individuals with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(2), 189–201.

Turner, M. A. (1999b). Generating novel ideas: fluency performance in high-functioning and learning disabled individuals with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *40*(2), 189–201.

Uddin, L. Q. (2015). Idiosyncratic connectivity in autism: Developmental and anatomical considerations. *Trends in Neurosciences*.

Vannorsdall, T. D., Maroof, D. a, Gordon, B., & Schretlen, D. J. (2012). Ideational fluency as a domain of human cognition. *Neuropsychology*, *26*(3), 400–5.

Wallach, M. A., & Kogan, N. (1965). Modes of thinking in young children. In *Modes of thinking in young children*. New York, NY: Holt, Rinehart and Wilson.

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research : Official Journal of the International Society for Autism Research*, *2*(3), 138–47.

World Health Organisation. (2012). Icd 10. *Who*.

Wykes, T., Huddy, V., Cellard, C., McGurk, S. R., & Czobor, P. (2011). A meta-analysis of cognitive remediation for schizophrenia: Methodology and effect sizes. *American Journal of Psychiatry*, *168*(5), 472–485.

# Part 2: Empirical Paper

# Designing an ecologically valid measure of executive function in children with Autistic Spectrum Disorder

# 1 Abstract

**Aims:** Despite increasing evidence of executive impairment in young people with autism, current measures of executive function (EF) are not fit for purpose. This is evidenced by poor correlation between existing measures and the types of everyday functional impairments experienced by those with the disorder. We sought to design and pilot an ecologically valid measure specifically for children with autism using a function-led approach.

**Method:** Seven tasks were designed around common features and symptoms reported in children with autism. These tasks formed the Ecologically Valid Test of Executive Disorder (Eco-TED). The measure was administered to twenty participants with a diagnosis of autistic spectrum disorder and twenty typically developing controls aged between 8 and 13 years with normal intellectual functioning (IQ>70). This paper investigates the psychometric properties of four tasks; the Luria, Pattern Drawing, Alternating Sequence and Storytelling tasks.

**Results:** Children with autism made significantly more errors on the Alternating Sequence task than controls ($p = .024$, ES = 0.36), with more additions to the patterns ($p = <.001$, ES = 0.61). On the storytelling task, those with autism recalled significantly fewer number of events ($p = <.001$, ES = 0.66), made more confabulations ($p = .001$, ES = 0.57) and made unclear references to key characters ($p = .037$, ES = 0.35). Test-retest reliabilities ranged from $r = .017$ to .648.

**Conclusion:** The Alternating Sequence and Storytelling tasks showed promising construct and criterion validity, warranting further development and evaluation. To our knowledge, the Eco-TED is the first function-led measure developed specifically for children with autism and as such improves on verisimilitude of current measures.

# 2 Introduction

Autistic spectrum disorder (ASD) (herein referred to as 'autism') is a pervasive developmental disorder characterised by persistent deficits in social communication and interaction and restricted and repetitive patterns of behaviour, interests or activities (DSM-V: American Psychiatric Association, 2013). Children and adults with autism often demonstrate a lack of coherent, goal directed behaviour, particularly in novel situations or environments with limited social structure (Russell, 1997). Several theories for autism have been proposed including; central coherence, male brain theory, and theory of mind (Baron-Cohen, 2002; Baron-cohen, Leslie, & Frith, 1985; Frith, 2003). Yet none of these theories can fully account for the range of difficulties displayed in those with the disorder. More recently, evidence has been mounting for an executive dysfunction account of autism which, as some suggest, may bring together existing theories and explain some of the problems typically reported in those with autism including; inflexibility, lack of symbolic play and impairment in discourse (Pennington, Bruce et al., 1997).

Executive functions are "those skills necessary for purposeful, goal-directed activity" (Anderson, 1998). The term is used to describe a range of cognitive processes including planning, working memory, mental flexibility, response initiation, inhibition, impulse control and monitoring of action (Alvarez & Emory, 2006; Robinson, Goddard, Dritschel, Wisley, & Howlin, 2009). These processes are thought to be governed predominantly by the prefrontal regions (Elliott, 2003). Impairments in executive functioning, as is seen in those with frontal lobe lesions, can have a wide ranging effect on the control and regulation of behaviour (Robbins, 1997). By drawing comparisons between those with frontal lobe lesions and those with autism,

researchers have begun to understand autism as an executive disorder (Ozonoff, Pennington & Rogers, 1991).

This hypothesis is supported by a range of studies that have demonstrated marked executive impairments in those with autism; with and without the presence of a co-morbid learning disability (Hughes, Russell, & Robbins, 1994; Pennington & Ozonoff, 1996; Rumsey & Hamburger, 1988). These studies tend to rely on traditional neuropsychological measures and suggest impairments on tests of planning, mental flexibility, inhibition and generativity. These impairments are present when compared to both typically developing controls and individuals with other neurodevelopmental disorders (such as ADHD, Tourette's and Dyslexia) (see reviews by Hill, 2004a, 2004b).

Planning is typically studied using a version of the Towers test (Tower of London, Tower of Hanoi or Stockings of Cambridge) and results have tended to indicate that those with autism require significantly more moves and demonstrate more rule violations than controls (Robinson et al., 2009). In tasks of mental flexibility (such as the Wisconsin Card Sorting Task), individuals with autism have shown increased perseveration and difficulty adapting to new rules (Hill, 2004b). Generativity which is described as the ability to spontaneously generate novel ideas without excessive pauses or errors (Pastor-Cerezuela, Fernandez-Andres, Feo-Alvarez, & Gonzalez-Sala, 2016), seems also to be affected in those with Autism (see Thesis Part 1: Literature Review). Similarly, whilst inhibition is generally thought to be preserved in individuals with autism (Bishop & Norbury, 2005), some studies have found difficulties in inhibiting a prepotent response, mainly through the use of the 'Go/No go', Windows and detour-

reaching tasks (Hill, 2004b; Robinson et al., 2009; Towgood, Meuwese, Gilbert, Turner, & Burgess, 2009).

Despite seemingly strong evidence for executive impairment, there also remains a fair degree of debate amongst the literature. Some studies have failed to replicate previous findings or find evidence of impairment (Kenworthy, Yerys, Anthony, & Wallace, 2008; Minshew, Goldstein, Muenz, & Payton, 1992; Rajendran & Mitchell, 2007). Others report superiority in the performance of some autistic individuals on tasks requiring executive abilities (Happe, 1999). Some say that this is a consequence of the heterogeneity that exists amongst the autism phenotype which means it is impossible to use executive impairments as a marker for the disorder (Hill, 2004b). These findings might seem to pose serious challenges to the executive dysfunction hypothesis (Griffith, Pennington, Wehner, & Rogers, 1999). However, anecdotal evidence from parents and clinicians suggests that children and adults with autism demonstrate a range of impairments indicative of executive function deficits (Kenworthy et al., 2008). This raises the question of why these deficits are not being consistently identified through existing executive measures.

Historically neuropsychological testing has been concerned with clinical diagnosis. As a result traditional measures were often developed from experimental investigations using non-clinical samples and they were designed for very specific aspects of executive function (Shallice & Burgess, 1991). In recent years the focus of testing has shifted and is now more concerned with assessing the impact of deficits on an individual's everyday functioning, informing psychological formulation and predicting outcomes (Burgess et al., 2006; Chaytor & Schmitter-Edgecombe, 2003).

These changes have increased the need for measures that have good clinical utility in terms of both diagnosis and intervention.

Despite this the tests themselves have not changed leading to questions as to how well existing measures capture real-life deficits. An example of this is seen in the frontal lobe literature where individuals with frontal lobe damage have been shown to perform with no impairments on tests of executive function, yet show difficulties in completing everyday tasks (White, Burgess, & Hill, 2009). Similarly within the Autism literature, investigations of executive function have failed to show a convincing relationship between impairment and autistic symptomology (Liss et al., 2001). Burgess et al. (2006) argue that this is because neuropsychological tests have not been adapted as their function changes; and that many are not used for the purpose for which they were originally intended. Traditional measures are based on construct-level theories rather than on observations of the populations they seek to understand (Wilson, Evans, Emslie, Alderman, & Burgess, 1998). This means that they often lack representativeness and generalisability to adults and children with particular disorders (Burgess et al., 2006). Therefore, there is a growing argument for the development of more "bespoke" neuropsychological tests that possess increased ecological validity and are designed with specific populations in mind (Burgess et al., 2006: p.194; Kenworthy et al., 2008).

In the context of neuropsychology, ecological validity describes the degree to which results obtained in controlled experimental conditions are representative of those that would be obtained in a naturalistic environment (Kvavilashvili & Ellis, 2004). Chaytor and Schmitter-Edgecombe (2003) make the point that diagnostic

validity is not necessarily synonymous with ecological validity. This is apparent in the autism literature, where the classic tests of executive function do not correlate well with autistic symptamology. One such example of this is a study into cognitive flexibility which found a marked difference between behavioural flexibility in the day-to-day environment and performance on cognitive flexibility tasks in those with autism (Geurts, Corbett, & Solomon, 2009). Therefore the current direction in neuropsychological research is the development of tests of executive function that translate into real-world settings, as it is this that we are most concerned about in clinical practice (Burgess et al., 2006).

In order for a test to posess ecological validity it needs to demonstrate both verisimilitude and veridicality (Kenworthy et al., 2008). The versimilitude of a test is the degree to which it acurately resembles the cognitive demands that exist in the everyday environment (Franzen and Willhelm, 1996). Veridicality is the degree to which test performance predicts everyday functioning (Gioia & Isquith, 2004). One criticism of classic neuropsychological tests is that they often lack the complexity that would exist in the real-world, where individuals are required to integrate multi-dimensional information (Wilson et al., 1998). Tests performed in clinic are often done so in a quiet, structured and unnatural environment far removed from the numerous extraneous variables that would be found in a real-word context (Shallice & Burgess, 1991; Teunisse et al., 2012). Hill (2004a) argues that what is needed are more naturalistic tests that build a stronger case for Executive Dysfunction in Autism. These tests should be developed with the autism phenotype in mind, rather than being borne out of brain injury research or by drawing comparisons with other clinical populations

such as those with Dysexecutive Syndrome (Ozonoff et al., 1991; Rajendran & Mitchell, 2007).

In order to make this shift, a different methodological stance is warranted. Moving from a top-down, construct-driven approach to one that is function-led, is an arguably more ecologically valid means of task development (Burgess et al., 2006). Tests developed in this way have already been shown to be more closely related to observed everyday symptoms than traditional tests (Wilson et al., 1998). Furthermore, in contrast to the argument that such tasks will be psychometrically unsound, many have been found to demonstrate good clinical utility (Burgess et al., 2006).

Currently the most ecologically valid performance measure of executive function available is the Behavioural Assessment of Dysexecutive Syndrome (BADS or BADS-C for children) (Emslie, Wilson, Burden, Nimmo-Smith, & Wilson, 2003; Wilson et al., 1998). This measure has good face vallidity and demonstrates increased verisimilitude. The BADS includes open-ended tasks that attempt to predict real-life problems (Kenworthy et al., 2008; Wilson et al., 1998). Unlike the classic tests, the BADS has also shown good correlation with autistic symptamology (Hill & Bird, 2006).

However depsite it's apparent utility and ecological validity, the BADS has demonstrated inadequate reliability (Henry & Bettenay, 2010). This is even when taking into account that test-retest reliabilities are likely to be lower in tests with increased verisimilitude, as individual's learn to adapt to the tasks presented (Chaytor & Schmitter-Edgecombe, 2003). Kenworthy et al. (2008) also state that the BADS does not possess veridicality and some have shown that the sub-tests do not correlate

with measures of everyday functioning in a brain injured sample (Wood & Liossi, 2006). In fact, it could be argued that the much of the success of the BADS relies upon the Dysexecutive Questionnaire (a 20 item parent-report questionnaire) that is administered alongside the sub-tests. This questionnaire has been shown to have greater sensitivity in measuring executive function in adolescents with autism (indicated by larger effect sizes) than the novel problem solving taks themselves (Channon, Charman, Heap, Crawford, & Rios, 2001).

In a review of executive function tests currently available, Henry and Bettenay (2010) conclude that no one test battery, including the BADS, assesses all five areas of executive function comprehensively. These five areas include; executive loaded working memory, fluency/reconstitution, inhibition, set shifting/switching and planning/problem-solving. Further, as with other tests of executive function, the BADS presents difficulties when trying to separate verbal and visuospatial skills for comparison making it difficult to identify the actual executive functions being measured (Henry & Bettenay, 2010).

In response to the above, we set out to develop and pilot an ecologically valid test of executive dysfunction specifically designed for children with Autism (the Eco-TED). We based the measure on the observed interaction between a child with autism and their environment, or in other words at the functional level of analysis by basing it on those impairments typically reported by parents and care-givers (Burgess et al., 2006). In using a function-led approach to task development, it was hypothesised that the measure would possess better veridicality and versimiltude and therefore be more highly correlated with measures of adaptive functioning. As such it would add to our

current understanding of executive function in those diagnosed with autism and have real-life application.

This research set out with the following aims:

1 To develop an ecologically valid measure of executive function that posesses both versimilitude and veridicalty

2 To conduct an initial evaluation of the psychometric properties of the measure including:

(a) Test-retest reliability

(b) Criterion validity

(c) Construct validity

# 3 Methods

## 3.1 Study Design

This study is a psychometric study that describes the development and piloting of the Eco-TED; a battery of seven tasks designed using a function-led approach for children with autism. This thesis focuses on the development and refinement of four of the seven tasks, with the remaining tasks described elsewhere (Bristow, 2016). The measure was administered to a clinical sample of children with autism and a matched control group in order to investigate its psychometric properties including criterion and construct validity. Test-retest reliability was assessed using a sub-sample of clinical and control participants.

## 3.2 Participants

Those with autism were recruited from a specialist social communication disorder unit and consisted of 20 children aged 8 to 12 years at the time of recruitment

(M = 135.28 months, SD = 3.25). The participants all had a previous diagnosis of either High Functioning Autism, Asperger's Syndrome or Autistic Spectrum Disorder based on the clinical consensus of a team of clinicians trained specifically in the assessment of social communication disorders. Diagnosis was based on information gathered from a range of sources including; The Developmental, Dimensional and Diagnostic Interview (3di) (Skuse et al., 2004), the Autism Diagnostic Observation Scale (ADOS) (Lord et al., 1989), school reports and clinical observations. Of the 20 clinical participants 10 were male and 10 Female. Six of the autistic group had a co-morbid diagnosis which included; Attention Deficit Hyperactivity Disorder (ADHD) (2 children), Attention Deficit Disorder (ADD) (1 child), Generalised Anxiety Disorder (1 child), Oppositional Defiant Disorder (ODD) (1 child); and Obsessive Compulsive Disorder (1 child).

The control sample consisted of 20 typically developing (TD) children aged 8 to 12 at time of recruitment (M = 131.05 months, SD = 14.85). There were 17 males and 3 females. Fifteen of these were recruited through a local school where the researcher had existing contacts, whilst five participants were recruited through convenience and snowball sampling. To ensure that the control participants did not present with autistic symptomology the Social Communications Disorder Checklist was completed for each participant. Children with a pre-existing diagnosis of a neurodevelopmental disorder were excluded from the control group. One of the control group had a co-morbid diagnosis of Ehlers-Danlos Syndrome (Type 3), three of the children had siblings with a diagnosis of autism and one of the control group had a sibling with a significant speech and language disorder.

For both the clinical and control samples, only children that had a full-scale intelligence quotient (FSIQ) within the normal range (FSIQ > 70) were included in the study. All participants spoke English as their first language. See Table 2.1 for a summary of participant characteristics.

Table 2.1. Participant characteristics for the ASD and TD group.

| | TD (n=20) *Mean (SD)* | ASD (n=20) *Mean (SD)* | *p* |
|---|---|---|---|
| Age in months | 131.05 (14.85) | 135.28 (3.25) | *ns* |
| FSIQ[a] | 107.60 (13.98) | 102.06 (14.43) | *ns* |
| Gender *(m:f)* | 17:3 | 10:10 | |
| Ethnicity *(n)* | | | |
| White British | 17 | 17 | |
| White Other | 2 | 0 | |
| Mixed Ethnicity | 1 | 3 | |
| Clinical Diagnosis *(n)* | | | |
| ASD[b] | | 11 | |
| HFA[c] | | 1 | |
| Asperger syndrome | | 8 | |

[a]FSIQ not available for two participants in the ASD group due to refusal to complete the test
[b]Autistic Spectrum Disorder
[c]High Functioning Autism

### 3.3 Design of the Eco-TED

The main aim of the project was to develop a function-led measure based on the difficulties most commonly reported by parents of children with autism. To do this we examined a large sample of data from the administration of the 3Di. The 3Di is a standardised interview that features a range of questions that can assist in the diagnosis

of ASD and/or other disorders (Skuse, 2013). The data was collected at the Social Communications Disorder Unit over a ten-year period and included parent responses for a large sample of children with autism, a non-ASD clinical group and a non-clinical group. Items pertaining to executive function (n = 33) were examined to distinguish which were more frequently rated as impaired in the ASD children compared with the non-ASD samples. Eleven items were identified and considered further (see Table 2.2).

Table 2.2. 3di items more frequently reported as impaired in children with ASD

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

*Items that formed the basis for tasks discussed in this paper.

### *3.3.1 Task Refinement*

The research team, which comprised Professor Paul Burgess and Dr Will Mandy (both of whom have experience in developing neuropsychological measures) and two doctoral students, engaged in an iterative process of task design. Following group discussion and initial generation of ideas, items were split equally amongst the doctoral students. Some items naturally grouped together (e.g. items (iii), (iv) and (v) were hypothesised to involve similar cognitive processes) and so were assigned to the same student.

Each student took responsibility for proposing four to five rudimentary tasks pertaining to those items. For example, item (ii) led to the generation of five task ideas (paper weaving, knot tying, friendship patterns, hand movement tasks and shoelace pattern generation). Following group discussion, some of the tasks were eliminated based on concerns such as children having differential experience of the tasks in everyday life (e.g. children that attend certain clubs might have had more practice of knot tying). The aim at this point was to keep the tasks analogous with the everyday activities captured by the 3di items, whilst also having clinical utility i.e. being quick and easy to administer. Proposed tasks were either based on entirely novel ideas or involved the identification of existing, non-copyrighted tasks that pertained to the 3Di items and could be adapted for use in a young autistic population. Those tasks that were considered most feasible were then taken forward and prototypes were constructed. These were subjected to a process of informal piloting and subsequent revisions (around three to five revisions including some major) for each task prototype. Piloting at this stage included video-feedback which assisted the team in deciding whether the task would be feasible to administer in clinic.

In addition to task construction, this process of piloting also led to the development of a task script which underwent a series of revisions (seven in total) (Appendix II). The task script was modelled on existing neuropsychological measures in terms of the level of detail included in instructions, discontinuation rules and directions for scoring. This was to ensure standardisation of administration.

The outcome of this process was a total of seven tasks, four of which are described in this thesis. The three remaining tasks were developed and refined by another doctoral trainee (Bristow, 2016). Development of the tasks from initial conception to agreement on the finalised tasks took twelve months. Details of trainee's individual contributions can be found in Appendix III.

## 3.4 Measures

### *3.4.1 Eco-Ted Tasks*

#### *3.4.1.1 Luria Test*

The 3di data indicated that children and adolescents with autism found it more difficult than their non-ASD peers to master the skill of tying their shoelaces. We hypothesised that there were several cognitive operations that may be involved with the execution of this task, namely the learning of multistep manualised sequences, motor programming/planning and coordination. Given that children are taught to tie their shoes by someone else, there is also the social aspect of imitation/copying another. The Luria hand movement test (also known as the fist-edge-palm test) seemed to incorporate all the above elements as well as being quick and easy to administer.

The test was originally devised as an assessment for individuals with brain pathology with impaired performance on the task noted in adults with dementia and Huntington's disease (Luria, 1980; Moses Jr., Golden, Berger, & Wisniewski, 1981;

Weiner, Hynan, Rossetti, & Falkowski, 2011). A version of the test is included in the Luria-Nebraska Neuropsychological Battery-Children's revision (Golden, 1991). This battery has demonstrated variable reliability and validity but has been commonly used in the assessment of children (Leark, 2004). The Luria hand movement test has been found to be sensitive to motor processing, spatial disorganisation and compulsiveness in children and adolescents (Davis, 2011). In a study comparing children with benign focal epilepsy and normal controls, the former demonstrated impaired sequential motor actions (Miziara, Giraldes de Manreza, Mansur, Conti Reed, & Buchpiguel, 2013). Similarly, studies employing the test have found motor impairment in children with brain damage and learning disability (Roy, Bottos, Pryde, & Dewey, 2004). These impairments include difficulties inhibiting responses when asked to repeat motor sequences a set number of times (Diamond, 2001). The fact that impairments on the task have been demonstrated in other populations with frontal lobe deficits suggested that it might have clinical utility in an autistic population.

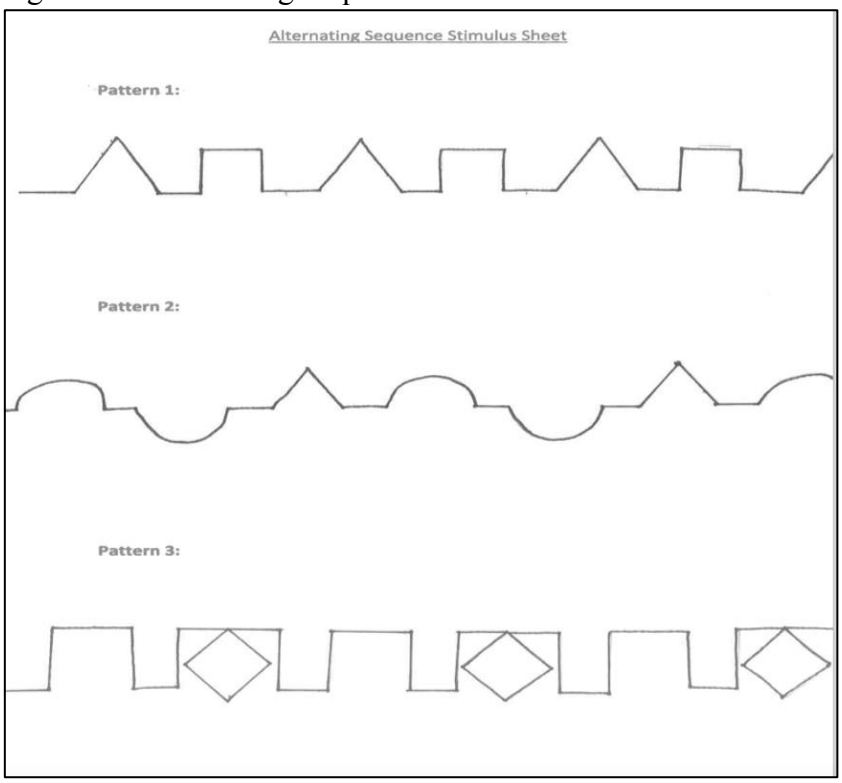
This subtest requires the child to copy a sequence of hand movements demonstrated by the researcher using their dominant hand. Initially the child is asked to touch the tip of their fingers in sequence using their thumb. This trial is to ensure that the child has the motor coordination abilities to continue with the subtest. Following this the researcher demonstrates a sequence of hand movements involving the 'fist', 'edge' and 'palm' of the hand in various orders, repeating the movement three times. The child is required to watch and then copy the sequence including the number of repetitions. There are nine trials. One point is awarded if the child gets the sequence correct and another point for the correct number of repetitions yielding an overall 'Total Score' out of a possible 18 points.

*3.4.1.2 Alternating Sequence Task*

This task was the second to be developed around the shoelaces item. Like the Luria test it is influenced by a measure that has previously been used in the assessment of complex cognition (see Arciniegas & Beresford, 2001). It involves the learning and execution of a sequence but requires a paper-and-pencil solution. The child is shown a repeated pattern on a piece of paper that goes from left to right of the page (see Figure 2.1). The researcher demonstrates how the pattern is drawn by tracing over the stimulus. The child is then asked to copy the pattern exactly as they see it without removing their pen from the paper. In demonstrating to the child how they should go about drawing the pattern, the researcher is mimicking real-life demands of the child watching and copying a caregiver when learning to tie shoes. Further, by completing the pattern without removing their hand from the paper they are having to follow rules whilst employing fine motor control as they would in the real-life scenario.

We developed the task further than that used previously by including two additional trials of increasing complexity. This was to reflect the fact that children are often taught increasingly more complex ways of tying their shoes. Time taken to complete the trial was recorded. Also, number of omissions, number of additions (additional shapes /lines not in original pattern), number of times the child retraced a line, number of errors (a composite of all three) and the length of the pattern drawn. The scoring was devised in this way so that there were variables capturing motor sequence learning, the child's ability to follow rules and their tendency to perseverate.

Figure 2.1. Alternating Sequence Task Stimulus




*3.4.1.3 Pattern Drawing Task*

This was the final task to be developed based on the 3di shoelace item. Unlike the previous two, this task was not developed from an existing measure. Rather, the idea came about from discussing the types of skills that are required by children when learning to tie their shoes; and from other activities a child might perform that uses similar skills. The task evolved from initial ideas of paper weaving and knot tying which, for various reasons touched upon earlier, were not practically feasible in our test battery. So, after several iterations the pattern drawing task was proposed which was thought to capture the key skills required to complete the original 3Di item but could also be administered easily in session.

The child is shown a picture of a shape and told that they will need to copy it. They are then provided with a response booklet and a demonstration is completed. The

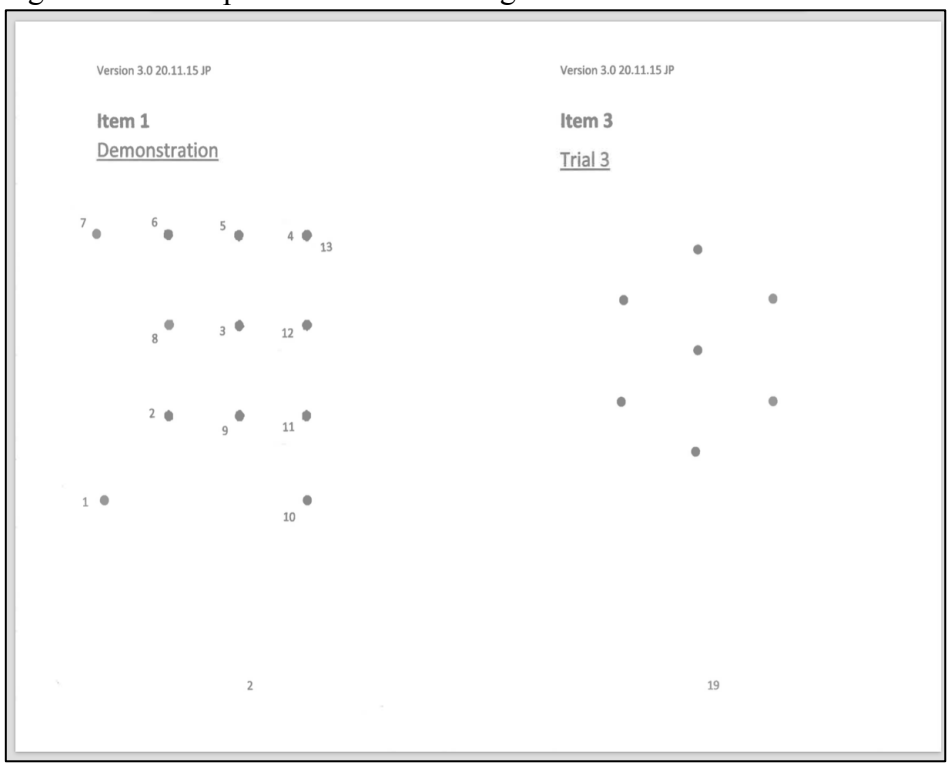
demonstration shows a series of numbered dots which the child is required to join in a sequential order, without removing pen from paper, to form the shape that they were originally shown. Once the child has completed the demonstration correctly, they advance to the trial proper. They are now asked to draw the shape again by joining the dots but this time there are no numbers to help them and they must remember how they drew the shape in the demonstration trial (see Figure 2.2 for example). They are reminded to keep pen to paper at all times and to draw the shape as quickly as possible. If the child draws the shape correctly, the researcher continues to the next item. There are three items in total, with the patterns increasing in complexity for each. If the trial is completed incorrectly the child is given two more chances (so a total of three trials per item), including two more demonstration items.

Children being taught to tie shoe laces would be given more than one opportunity to learn the skill and more than one demonstration following an unsuccessful attempt, hence the inclusion of several demonstrations and trials per item. However, important to note is that the learning aspect of this task is through the child joining the dots to form the shape rather than by copying the researcher. This was a deliberate attempt to remove the influence of the researcher unlike in the previous tasks. It has been shown elsewhere that social cognition can impact on an autistic child's performance on researcher-administered tasks (Kenworthy et al., 2008; Ozonoff, 1995), therefore we wanted one of the tasks to control for this confounder.

For this subtest, scoring is based on the time taken to complete the pattern for each trial and the number of errors. The latter variable is calculated by subtracting the number of correctly connected dots from the total number of dots that should be connected to complete the pattern (*Total no. correct connections – no. completed*

*correctly = error*). So, for item one, the child is required to make 12 connections to draw the pattern correctly. If they connected the first four dots but then made a mistake or failed to complete the trial, the number of errors would be 8 (12-4 = 8). From these raw composite scores for total time and total number of errors are calculated.

Figure 2.2. Example of Pattern Drawing Task Stimulus




*3.4.1.4 Storytelling Task*

This task was developed in response to the 3di items: *Can X give an easy to follow account of past events such as a birthday party?* and *What about difficulty in telling a story or describing what* they *have done*? These items are likely to draw on a range of linguistic, cognitive and socio-cognitive abilities but were generally thought to pertain to the cognitive operations of narrative coherence and working memory. Like the pattern drawing task, this task arose from a purely iterative process of task

design rather than being based upon existing measures. There are three parts to this subtest.

In the first part the child is played an audio-recording of a short story titled 'Lunchtime'. They are instructed to listen carefully as they will be asked to re-tell the story once it has finished. In addition, they are provided with a cartoon strip depicting what happens which they may use as a visual prompt. Once the recording is finished, the child is then asked to re-tell the story in as much detail as possible. Their response is audio recorded for ease of scoring. In the second part of the task, the above procedure is repeated but this time the child listens to a different story; 'Doris the Cat'. For this task, the child is shown picture prompt cards depicting key elements of the story (there are four in total). Again, they are asked to recall the story in as much detail as possible but this time they do not have the prompt cards as a visual prompt during recall.

The main considerations when designing this subtest were that the stories needed to be standardised, brief to administer but also detailed enough to avoid a ceiling effect. A range of stories were devised by the researcher which took into consideration the reading age of the intended population (8-12 years). These stories were standardised by ensuring that they all contained the same number of events and action frames. The research team then reached a consensus on which of the stories were suitable for inclusion and the two detailed above were selected. By including a visual component as well as an auditory component, the tasks accessed both verbal and visual domains as would such a task if performed in real-life.

The final part of the subtest was performed at the end of the testing session, once the child had completed all subsequent neuropsychological measures. They were instructed "*Now I want you to tell me everything that has happened since you got*

*here"*. This was directly comparable to the 3di item in question and resembled an everyday task that a child might encounter. Further, because the researcher had administered the testing session then this task was inherently standardised as the researcher knew exactly what the child's account should contain.

Many variables were produced during the scoring of this subtest which is indicative of the complexity of examining narrative coherence / recall (e.g. Baesler, 1995). Recordings were rated after the testing session and a series of composite scores across all three stories created (Appendix IV):

- Time taken to recall the story
- Number of events recalled
- Number of errors in order of events (e.g. how many times the child got the sequence of events wrong)
- Number of prompts required to continue telling the story
- Number of pauses over 2 seconds
- Number of Theory of Mind (ToM) or mentalising words in account
- Number of new elements (confabulations) – these were required to be gross additions or substitutions such as recalling a dog instead of a cat
- Number of repetitions (how many times the child repeats an element of the story)
- Number of incomplete sentences
- Number of times child referred to a character or person without it being clear who they were referring to

#### *3.4.1.5 Scoring and Key Outcome Measures*

For each Eco-TED measure, a primary outcome variable(s) was identified. This was an a prior outcome variable that we hypothesised would give us the most relevant information about task performance (see Table 2.3):

Table 2.3 A priori key outcome variables for each task

| Measure | Key Outcome Variable(s) |
|---|---|
| Luria Task | Total Score |
| Alternating Sequence Task | Total Errors |
| Pattern Drawing Task | Total Errors |
| Storytelling Task | Total Number of Events |
| | Total Errors |

### *3.4.2 Existing Neuropsychological Measures*

As a means of measuring the criterion validity of the Eco-TED, two-subtests of the BADS-C (Emslie et al., 2003) were also included in the assessment battery . These comprised the Zoo Map and Six Part test. The BADS-C was chosen for comparison as it is currently the most ecologically valid measure of executive function available and has been demonstrated to have good construct validity (Baron, 2007; Engel-Yeger, Josman, & Rosenblum, 2009; Roy, Allain, Roulin, Fournet, & Le Gall, 2015) Furthermore, the chosen subtests have been shown to discriminate between adults with ASD and TD controls (Hill & Bird, 2006).

The Zoo Map test is a paper-and-pencil test in which the child is presented with a map of an imaginary zoo and asked to plan a route around it. There are two parts to the task. The first part is minimally structured and relies predominantly on the child's

ability to follow instructions without error. The second part of the subtest relies more heavily on the child's planning abilities. Performance is assessed based on accuracy and time taken to complete trials.

The Six Part Test relies on executive abilities including planning, task scheduling and performance monitoring. The child is given five minutes to cover six different tasks. They are not expected to finish the tasks but must follow certain rules such as not completing two parts of the same task consecutively. To do well on this subtest the child must be able to generate strategies. Performance is assessed based on the number of tasks completed, number of times rules are broken and amount of time spent on each task.

### *3.4.3 Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II) (Weschler, 1999)*

As participants were required to have a FSIQ estimate within the normal range, a measure of intellectual functioning was also included in the assessment battery. Five participants in the clinical sample had received an IQ estimate at the time of diagnosis. For all remaining participants, the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II) (Weschler, 1999) was completed. This measure is favoured as a research tool given its brevity. It has been shown to have acceptable reliability and validity and is highly correlated with the full scale (Homack & Reynolds, 2007; Irby & Floyd, 2013; McCrimmon & Smith, 2012).

### *3.4.4 Parent Report Questionnaires*

A series of validated parent report measures were used to assess characteristic of our samples and to provide an additional means of assessing criterion validity.

#### *3.4.4.1 Social Communication Disorders Checklist (SCDC)*

The SCDC (Skuse et al., 1997) is a 12-item screening questionnaire for autistic traits. The measure has been shown to have excellent internal consistency and good reliability, accurately discriminating those with ASD from non-clinical samples (Skuse, Mandy, & Scourfield, 2005). The SCDC was used in the current study as an assessment of autistic symptomology in the clinical sample and as a means of screening the control participants for autistic feature. A score of nine points or above is suggestive of autism (Wilkinson, 2010).

#### *3.4.4.2 Behaviour Rating Inventory of Executive Function (BRIEF®)*

The BRIEF® (Gioia, Isquith, Guy, & Kenworthy, 2000) is an 86 item questionnaire designed to assess executive function in the home and school environment. These items pertain to eight clinical scales; Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organisation of Materials and Monitoring. From these an overall score (the Global Executive Composite) can be calculated. In the current study, parent-report forms were used to gain an idea of our samples overall executive function characteristics. These provided an additional measure against which to assess the criterion validity of the Eco-TED and to assess how well it correlates with everyday outcome variables (its veridicality).

#### *3.4.4.3 Strengths and Difficulties Questionnaire (SDQ)*

The SDQ is a brief 25-item behavioural screening questionnaire designed for children aged 3-16 years (Youth in Mind, n.d.). The current study used the parent-report version which asks parents to rate symptoms based on five areas of functioning; emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour. The first four are totalled to gain an overall 'total

difficulties score'. In addition, there is an impact supplement which indicates chronicity, distress and social impairment of any difficulties. Scores on the SDQ were used to characterise the clinical and control samples (Appendix V).

### 3.5 Procedure

Ethical approval was granted by Westminster NHS Research Ethics Committee (ref: 15/LO/1332). Permission to recruit the clinical sample was also obtained from the Joint Research and Development Office of the Foundation Trust from which the participants were recruited (Appendix VI). Copies of the participant information sheets and consent forms are included in Appendices VII-XI).

The ASD population were approached by a member of the clinical team if they had previously given permission to be contacted for research purposes. For the control sample, parents were sent information sheets by the school and asked to opt in to the study (Appendix XII). Assent was obtained from all children prior to participation and consent was sought from parents or primary caregivers. Parents were also asked to complete three parent-report questionnaires. The clinical group were tested at University College London or at the participant's home depending on which was most convenient for the family. The control group were tested at their school except for a small number of children that comprised the convenience sample who were tested at home. All testing sessions took place in a quiet room, free of distractions and lasted approximately 90 minutes. The research team consisted of two researchers and three research assistants who had all been trained in task administration to ensure consistency in data collection. Only the participant and a maximum of two researchers were present whilst the session was in progress.

The assessment battery consisted of the Eco-TED and the additional neuropsychological measures detailed above. Participants were offered frequent breaks and refreshments to combat testing fatigue. They were given a £5 voucher as a reward for taking part and were entered into a prize draw to win a further £50 voucher. Twenty participants were approached to take part in the study a second time (10 clinical and 10 control) one month after their initial testing session. This allowed calculation of test-retest reliabilities. Participants were offered an additional £5 voucher for their participation. Of these, six participants declined or did not respond meaning that six clinical and eight control participants were re-tested.

## 3.6 Statistical Analysis

### *3.6.1 Assessing Reliability and Validity*

Given that the primary remit of this study was to develop a new measure of executive function and assess its psychometric properties, the initial stage of analysis compared scores for the Eco-Ted tasks across two time points to determine test-retest reliabilities. Next, we conducted a group difference analysis on the raw scores and composite raw scores for the clinical and control groups on each of the Eco-Ted tasks. This was to determine whether any of the tasks could differentiate between the two groups, in essence assessing its construct validity. Group comparisons were also conducted for scores on the SCDC and SDQ. Finally, concurrent criterion validity was assessed by measuring correlations of the tasks with existing 'gold standard' measures of executive function, namely the BADS-C and the BRIEF.

Where data was not normally distributed, non-parametric tests were applied rather than transforming the data, given that this method is arguably more robust (Field, 2013; Glass, Peckham, & Sanders, 1972). Correlations were conducted using

Kendall's-tau as this statistic is a better option for smaller data sets with tied ranks (Field, 2013). Where Independent t-tests were conducted, adjusted t statistics are reported if Levene's test for homogeneity of variance were significant. Effect size calculations are described according to the conventions outlined by Cohen (1992).

### *3.6.2 Dealing with Outliers and Multiplicity*

The data did contain some outliers (12 data points in all), however the decision was taken not to trim or manipulate these. Given the heterogeneity that is often displayed amongst those with autism, it was felt that these outliers might reflect actual performance rather than be a consequence of error. Thus, modifying the data would have been over-rigorous (Tukey, 1960).

As the analyses took an exploratory stance, multiple tests were run which increased the chance of multiplicity and type I error. However, this exploratory approach is not unusual (Barker, Pistrang, & Elliott, 2002) and was necessary for us to better understand the variables and any group differences captured by our measure. Given the chance of type I error, we ensured that any interpretations of the data remained tentative. A statistical correction such as the Bonferroni was not applied given that it would have been too conservative (Abdi, 2007) and increased the chance of type II error.

### *3.6.3 Multiple-Case Series Analysis*

It is argued that a key characteristic of the autistic phenotype is the large degree of heterogeneity amongst those with the disorder (Hill & Bird, 2006). By looking at performance using group means we risked creating an "averaging artefact" (Shallice & Evans, 1978), whereby those autistic participants that showed more extreme variability in their performance (both impairments and supra-normal abilities) might

be missed. We therefore adopted the multiple-case series approach to compare within-subject performance. This method is described in Towgood et al. (2009) and gave an indication of how well our measure captured within-subject heterogeneity, which we examined in two ways.

Firstly, we created normative z-scores which are scores based on the performance of the matched control group. This method is one that is often used in neuropsychology to look at within-subject variability, and treats the control group as a 'normative reference sample' (Crawford & Garthwaite, 2005). Z-scores were reversed for those measures where high scores indicated greater impairment. This was so that a lower score on all measures pertained to a greater deficit and reduced the complexity of the data for greater parsimony (Anglim, 2009).

Z-scores were calculated as outlined by Owen (1985):

$$Z = (X_i - X_{ref}) / s_{ref}$$

(where: $X_i$ = individualised raw Score; $X_{ref}$ = reference mean score and $s_{ref}$ = reference standard deviation).

We used these normative z-scores to look at individual profiles of performance graphically. Next, we calculated a within-subject standard deviation of the mean z scores across all tasks to see whether those with autism showed a greater variation in performance (as indicated by a higher standard deviation of scores).

# 4 Results

## 4.1 Sample Characteristics: SCDC and SDQ

As expected, those with ASD scored significantly higher (M = 16.65, SD = 4.84) on the SCDC than controls (M = 1.5, SD = 1.85; U = 1.50, p = <.01). However, two of the clinical participants (participant 30 and 40 with a diagnosis of Asperger's and

ASD respectively) had an SCDC score that was below the suggested cut-off score for autism (n = 9). Despite this they were not identified as outliers as per their scores on all other measures relative to the other clinical participants. Given that they had received a diagnosis of Autism previously, their data was retained for analyses. None of the control groups scored above the cut-off on the SCDC.

An independent samples t-test was conducted to compare the overall difficulties scores for the ASD and Control group on the SDQ. The total difficulties score was significantly higher for the ASD group (M= 18.80, SD= 6.86) than for the TD controls (M= 4.65, SD= 4.27; t (31.79) = -7.83, p<.001).

### 4.2 Test – Retest Reliability

Nonparametric correlation coefficients are displayed in Table 2.4 for each of the measures. Insufficient variation in participant scores meant that test-retest reliability could not be calculated for the number of incorrect responses on the Luria task and total omissions on the Alternating Sequence task. Although several of the correlations were significant for the Alternating Sequence Task, none reached the acceptable level of 0.7-0.9 that is commonly suggested for psychometric measures (Anastasi & Urbina, 1997).

Table 2.4 Test-Retest Reliability of Eco-Ted Sub-measures

| Task | Measure | Mean (SD) (Time 1) | Mean (SD) (Time 2) | Stability Coefficient *(τ)* | P |
|---|---|---|---|---|---|
| Luria | No. Correct | 7.93 (1.07) | 8.21 (1.42) | .066 | .786 |
| | No. Partially Corr. | 0.86 (1.03) | 0.79 (1.42) | .017 | .944 |
| | Total Score | 16.71 (1.27) | 17.21 (1.42) | .097 | .686 |
| PD Task | Total Time | 47.84 (17.25) | 42.73 (17.36) | -.026 | .903 |
| | Total Trials | 5.00 (3.96) | 4.00 (1.11) | -.089 | .723 |
| | Total Errors | 8.38 (6.87) | 4.14 (6.13) | .092 | .691 |
| AS Task | Total Time | 86.31 (25.52) | 79.34 (26.92) | .648 | .001* |
| | Total Length | 52.06 (6.28) | 55.15 (7.09) | .398 | .048* |
| | Total Traced | 2.50 (5.33) | 2.64 (6.32) | .556 | .018* |
| | Total Additions | 0.64 (0.84) | 0.36 (6.42) | .404 | .117 |
| | Total Errors | 3.14 (5.80) | 3.36 (6.42) | .556 | .013* |

| | | | | | |
|---|---|---|---|---|---|
| Storytelling | Proportion Events[a] | 0.59 | 0.57 | .208 | .316 |
| | | (0.15) | (0.17) | | |
| | Proportion Errors[a] | 0.39 | 0.29 | .235 | .314 |
| | | (0.33) | (0.22) | | |
| | Total Time | 313.43 | 291.79 | .165 | .412 |
| | | (88.94) | (67.32) | | |
| | Total Prompts | 0.43 | 0.57 | -.116 | .651 |
| | | (0.94) | (0.94) | | |
| | Total Pauses | 2.93 | 0.36 | -.203 | .394 |
| | | (3.00) | (0.63) | | |
| | Total ToM | 2.00 | 3.57 | .329 | .145 |
| | | (1.75) | (1.40) | | |
| | Total Confabs | 0.57 | 0.64 | -.132 | .603 |
| | | (1.40) | (0.93) | | |
| | Total Reps | 0.14 | 0.29 | .345 | .203 |
| | | (0.36) | (0.61) | | |
| | Total Incomplete | 0.71 | 0.79 | .397 | .111 |
| | Sentences | (0.99) | (1.37) | | |
| | Total Unclear | 1.21 | 2.50 | .083 | .721 |
| | Character | (1.19) | (1.65) | | |

AS Task: Alternating Sequence Task

PD Task: Pattern Drawing Task

[a]Proportion of Events and Proportion of Errors were compared between time one and time two. This is because in the second testing session only the Eco-TED was performed meaning that total scores were not directly comparable for this measure.

## 4.3 Assessing Validity

### *4.3.1 Criterion Validity: The BADS-C and BRIEF*

Correlational analyses were conducted between the subtests of the Eco-TED and; three BADS-C subtests and the BRIEF General Executive Composite. Significant correlations are shown in Table 2.5. As can be seen, only the Alternating Sequence Task and the Storytelling Task had composite measures that significantly correlated with the BRIEF, with the strongest correlation evident between the BRIEF GEC and the Total Additions composite score of the Alternating Sequence Task. There were no significant correlations found between the Six-Part test and the Eco-TED and only two significant correlations with the Zoo Map test (Alternating Sequence Task Total Additions and Storytelling Unclear Characters).

### *4.3.2 Veridicality: Correlation with a Measure of Symptom Severity*

There were significant correlations between SCDC total score and; the Alternating Sequence Task (Total Additions and Total Errors). Also between SCDC total score and the Storytelling Task (Number of Events, Number of Errors and Number of Confabulations). For the remaining tasks, no other findings were significant. Correlational analyses were also conducted with the 'Impact Score' of the SDQ. There were significant correlations with all measures named above except for Total Confabulations and Total Unclear Characters. In addition, SDQ Impact Score was also correlated with Total Omissions on the Alternating Sequence Task (see Table 2.5).

Table 2.5 Correlations between Eco-Ted and Existing Measures of Executive Function and Autistic Symptomology

| Task | Measure | BRIEF GEC | BADS-C Zoo Map 1 | BADS-C Zoo Map 2 | BADS-C Six-part | SCDC Total Score | SDQ Impact Score |
|---|---|---|---|---|---|---|---|
| Luria | Total Score[a] | .011 | .196 | .168 | -.145 | -.043 | .019 |
| Pattern Drawing | Total Time | .203 | -.207 | -.129 | .035 | .182 | .189 |
| | Total Errors[a] | -.182 | -.156 | -.008 | -.013 | -.095 | -.045 |
| AS Task | Total Additions | .502** | -.190 | -.368** | -.100 | .460** | .537** |
| | Total Omissions | .215 | -.175 | -.254 | .044 | .223 | .331* |
| | Total Errors | .234* | -.155 | -.256 | .013 | .291* | .273* |
| Storytelling | Total Events | -.290* | .040 | .152 | .109 | -.358** | -.459** |
| | Total Errors | -.324* | .050 | .206 | -.033 | -.386** | -.362** |
| | Total Confabs | .345* | -.036 | -.248 | -.152 | .357** | .282 |
| | Total Unc. Charac | .204 | -.297* | -.215 | .099 | .175 | .209 |

* P < .05
** P < 0.1
AS Task: Alternating Sequence Task

## 4.4 Construct Validity: Comparison of Scores on the Eco-TED measures

### *4.4.1 Luria Test*

There were no significant differences in Total Scores between the autism and control groups on the Luria Test. The means and standard deviations of the control group (M = 15.65, SD = 2.37) and the clinical group (M = 14.80, SD = 3.55) were close to the total available score (18) suggesting a ceiling effect.

### *4.4.2 Pattern Drawing Task*

The primary outcome measure for the Pattern Drawing Task was Total Number of Errors. On this measure those with autism actually made fewer errors compared with control participants although this finding was not significant (see Table 2.6). Only the raw composite score 'Total Time' differentiated between the two groups (U= 114.00, p = .033); with Autistic participants taking significantly longer to complete the task than controls. This gave rise to a medium sized effect.

### *4.4.3 Alternating Sequence Task*

There were significant group differences for three of the measures, including for the primary outcome measure; Total Errors. Both Total Time (t (37) = -2.519, p = .016, two-tailed) and Total Errors (U= 110.50, p = .024) demonstrated a medium sized effect. Those with autism took significantly longer to complete the task and made significantly more errors than controls. For Total Additions, which demonstrated a large effect, the Autism group made significantly more additions to the patterns (U= 66.00, p = <.001). Two outliers were identified for Total Additions and Total Time but their inclusion did not significantly affect the outcome and so were included in the final analysis (see Table 2.6).

Table 2.6 Mean group comparisons of clinical and control groups on the Luria, Pattern Drawing and Alternating Sequence Tasks

| Task | Measure | N | Mean Score Control (SD) | Mean Score Autism (SD) | P | Effect Size (*r*) |
|---|---|---|---|---|---|---|
| Luria | Total Score[a] | 40 | 15.65 (2.37) | 14.80 (3.55) | .631 | -0.08 |
| Pattern Drawing | Total Time | 39 | 61.03 (29.42) | 91.44 (57.99) | .033* | 0.34 |
| | Total Errors[a] | 39 | 13.60 (12.76) | 11.42 (12.48) | .498 | 0.11 |
| Alternating Sequence | Total Time | 39 | 77.37 (18.03) | 95.78 (26.57) | .016* | 0.38 |
| | Total Length | 39 | 52.80 (5.46) | 51.26 (7.36) | .694 | 0.06 |
| | Total Traced | 39 | 1.79 (4.37) | 1.95 (4.52) | .926 | 0.02 |
| | Total Omissions | 39 | 1.00 (2.40) | 2.50 (3.90) | .111 | 0.25 |
| | Total Additions | 39 | 0.21 (0.54) | 2.50 (3.49) | <.001 * | 0.61 |
| | Total Errors[a] | 39 | 3.53 (5.27) | 6.70 (6.57) | .024* | 0.36 |

[a]Primary outcome measures

### *4.4.4 Storytelling Task*

There were five composite measures on the Storytelling Task where performance of the clinical and control group differed significantly from one-another. This included for the two primary outcome measures; total number of events recalled and total number of errors made in the sequence of events. A large effect size was found for Total Number of Events with the autistic group recalling significantly fewer number of events in the story than controls ($t(34) = 5.071$, $p<001$). This meant that, overall, the autistic participants spent less time recalling the stories ($U = 99.50$, $p = .049$). Although those with autism made fewer errors in the sequence of events than controls ($t(34) = 2.891$, $p = .007$), they were significantly more likely to confabulate ($U = 76.50$, $p = .001$). Amongst the autistic group, there were also significantly more instances where the child's account was difficult to follow, as measured by the number of occasions where the researcher was not clear which character the child was talking about ($t(34) = -2.177$, $p = .037$) (Table 2.7).

Table 2.7 Mean group comparisons of clinical and control groups on the Storytelling Task

| Measure | N | Mean Score Control (SD) | Mean Score Autism (SD) | P | Effect Size (*r*) |
|---|---|---|---|---|---|
| Total No. of Events[a] | 36 | 29.76 (5.03) | 18.95 (7.39) | <.001* | 0.66 |
| Total No. of Errors[a] | 36 | 2.47 (1.63) | 1.11 (1.20) | .007* | 0.44 |
| Total Time | 36 | 371.47 (118.23) | 300.53 (138.86) | .049* | 0.33 |
| Total No. of Prompts | 36 | 0.24 (0.75) | 0.84 (1.26) | .083 | 0.29 |
| Total No. of Pauses | 36 | 4.65 (3.76) | 3.63 (4.56) | .133 | 0.25 |
| Total No. ToM | 36 | 2.41 (1.33) | 1.68 (1.60) | .077 | 0.29 |
| Total No. Confabs | 36 | .00 (.00) | 1.42 (1.74) | .001* | 0.57 |
| Total No. Repetitions | 36 | 0.29 (0.59) | 0.21 (0.54) | .581 | 0.09 |
| Total Incomplete Sent. | 36 | 0.88 (1.17) | 2.05 (2.86) | .241 | 0.20 |
| Total Unclear Character | 36 | 1.24 (1.20) | 2.32 (1.70) | .037* | 0.35 |

[a] Primary outcome measure

## 4.5 Multiple-Case Series Analysis (Within-Subject Variability)

As can be seen in Figures 2.3 and 2.4, there did seem to be some difference in variability in performance for the autistic participants compared with the control participants. Visual inspection suggested that there was greater heterogeneity in those with autism but in general differences were most apparent in relation to impaired performance with few obvious supernormal peaks amongst the autistic sample.

When we looked at within-subject variability using standard deviations of participant's z scores, the control group's standard deviations ranged from 0.42 to 1.51 (mean = 0.85). For the autistic participants, this was 0.49 to 1.99 (mean = 0.99), suggesting that there was slightly greater variability in individual performance for those with autism. However, this difference was not statistically significant ($U$ = 148.00, p = .160).

Figure 2.3 Graph showing within-subject variability for control participants

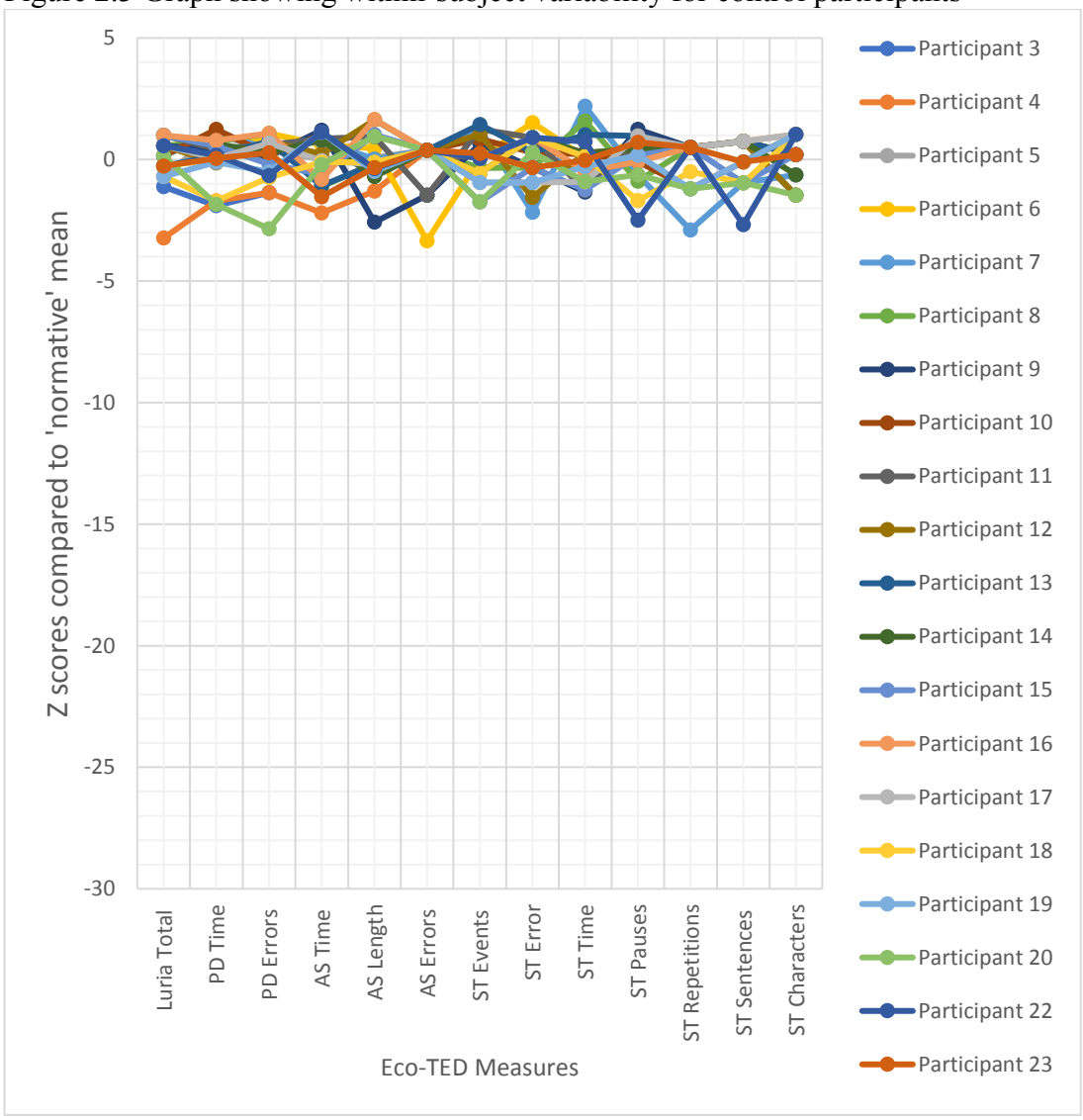
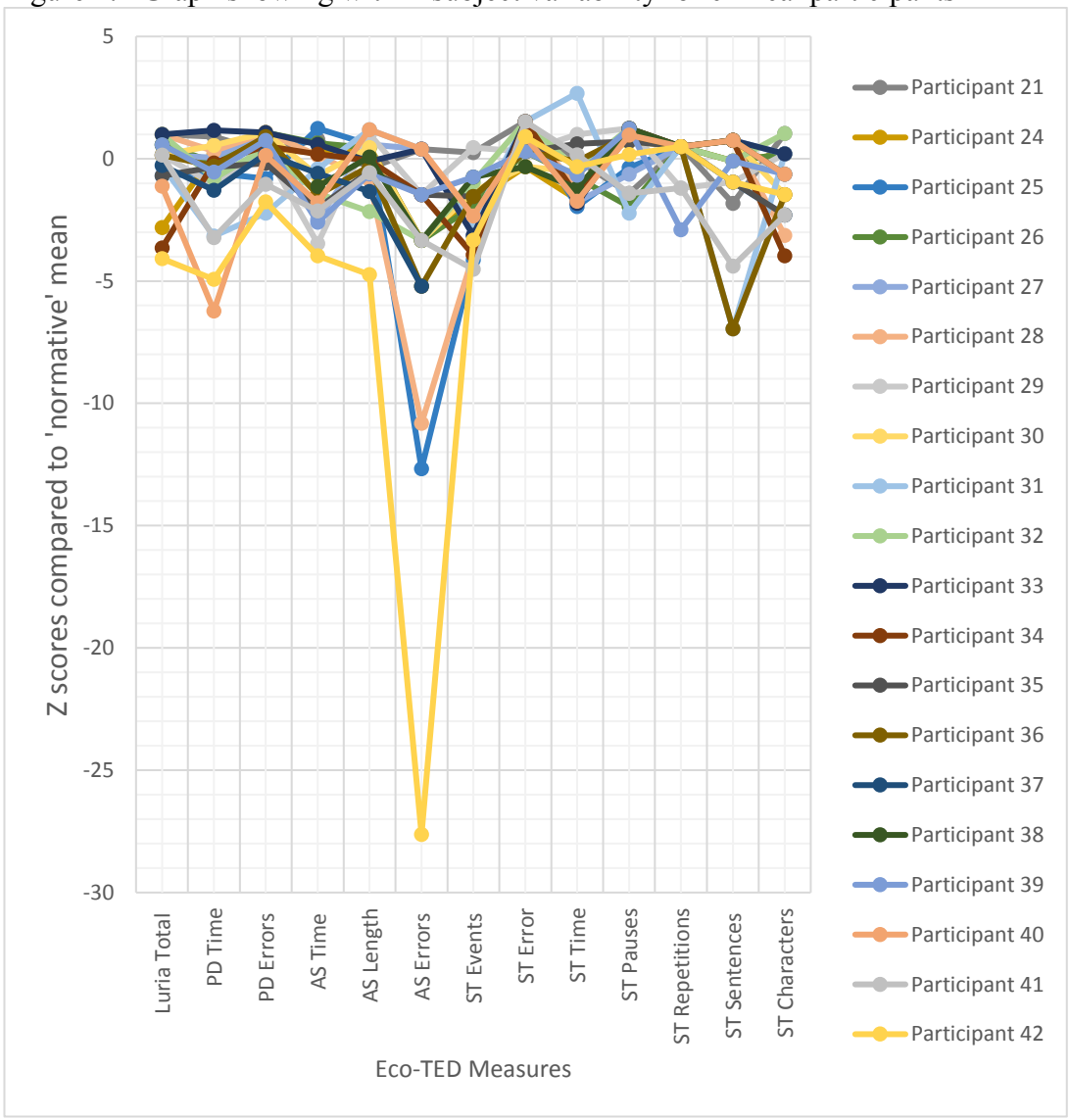


Figure 2.4 Graph showing within-subject variability for clinical participants



# 5 Discussion

The key aim of this study was to develop and pilot a new measure of executive function for children with autism (the Eco-TED). We did this using a bottom-up approach as advocated elsewhere (Burgess et al., 2006). This paper describes an evaluation of the psychometric properties for four of the seven Eco-TED tasks; including how well the tasks discriminated between those with and without autism. Of these four tasks (Luria Test, Alternating Sequence, Pattern Drawing and Storytelling Tasks) the latter two showed group differences that warrant further investigation through replication. Whilst the primary aim of the study was the development and initial psychometric evaluation of the tasks, we also speculate on what executive impairments may be involved.

## 5.1 Task Reliability

Although some of the test-retest reliabilities of the Eco-TED tasks were significant, none reached the acceptable level for psychometric measures (Anastasi & Urbina, 1997). Henry & Bettenay (2010) argue that this is inevitable when developing measures that have increased ecological validity. This is because such measures rely on novelty which is compromised when administering the measure a second time. This may have been particularly pertinent in our study as participants were re-tested just one month after the initial session. Although the test-retest reliabilities were low, they were in some cases similar to those achieved by other ecologically valid measures such as the BADS-C (Lezak, Howieson, Bigler, & Tranel, 2012). There were mixed results as to whether the children did better or worse across the tasks at the second administration, suggesting that poor test-retest reliability was not simply a result of practice effects as would be expected (Chaytor & Schmitter-Edgecombe, 2003).

Certainly, any future development of the Eco-TED would seek to improve on test-retest reliabilities as these are essential to any standardised measure (Anastasi & Urbina, 1997). Increasing the sample size, improving standardisation of the testing sessions and assessing inter-rater reliabilities might address some of the confounding factors that contributed towards poor test-retest reliabilities. These findings also raise a wider question of how to ensure the development of measures that are psychometrically sound yet retain ecological validity; a dilemma that is already debated within the literature (see Burgess et al., 2006).

## 5.2 Task Validity

### *5.2.1 Luria Test*

Comparison of performance by the clinical and control groups on the Luria test revealed no significant group differences. The measure was unable to differentiate between those with and without autism and thus demonstrated poor construct validity. Assessment of its criterion validity indicated no significant correlations with existing measures of executive function or with measures of symptom severity. This would suggest that the Luria test does not possess adequate construct or criterion validity to be included in further development of the measure.

These finding are interesting given that children with executive impairments caused by other frontal lobe conditions have shown poorer performance on the Luria test (Miziara et al., 2013). Impairment on this task is said to be a "soft-sign" in disorders associated with frontal lobe conditions including ADHD and Schizophrenia (Cobert, 2013, p.35). Therefore, one would have expected the autistic participants in our study to perform more poorly on this task, given proposed executive deficits. A further observation was that means for those with and without autism were close to the

overall score available for this task, suggesting a ceiling effect. A similar form of the task is included in the Luria-Nebraska Neuropsychological Battery for Children, which is specifically aimed at children aged 8-12 years (Leark, 2004) so it is surprising that there was a ceiling effect amongst our participants. These findings would seem to provide additional support for the view that current neuropsychological measures are not useful in assessing executive impairments in those with autism.

One explanation is that our use of the test was adapted so that it more accurately reflected the cognitive demands of the original 3Di item i.e. having to follow specific patterns whilst obeying instruction. Points were awarded to participants for replicating both the hand movements and the number of repetitions demonstrated by the examiner. In other versions of the test patients are asked to repeat hand movements continuously until the examiner says stop. It is possible that there may have been greater perseveration in our clinical sample if they were asked to repeat the movements for a longer period. In addition, the Luria test has been linked to more areas of the brain than just the frontal lobe. For example, activation of the cerebellum and parietal lobes have also been indicated through imaging studies (Umetsu et al., 2002). Therefore, it may be that the test is not sensitive and specific enough to capture the executive deficits present in those with autism.

### *5.2.2 Pattern Drawing Task*

As with the Luria test, the Pattern Drawing task lacked construct validity failing to find group differences between those with and without autism on the primary outcome variable ('total errors'). In addition, criterion validity for the task was poor with no significant correlations between the task and existing measures. Further, there were a large range in scores for this task in both the clinical and control samples

suggesting possible effects of outliers or skewed data. Contrary to our predictions, those with autism made fewer errors on this task than typically developing children. Although this finding was not significant especially when taking into consideration the range in scores, it is surprising given that children with autism have previously been shown to display impairments in procedural learning and motor sequencing (Gidley Larson & Mostofsky, 2008; Mostofsky, Goldberg, Landa, & Denckla, 2000). Again, these findings may suggest that the Pattern Drawing task is not sensitive enough to capture these types of executive impairments. However, administering the task to a larger sample might be warranted to discount the possibility that the task is able to identify the type of supra-normal abilities that some children with autism demonstrate (e.g. Towgood et al., 2009).

Performance on the Pattern Drawing task was not significantly correlated with scores on the SCDC and SDQ (both measures of functional symptoms). This would suggest that the Pattern Drawing task does not possess good veridicality (Kenworthy et al., 2008).

There was one significant group difference for the outcome variable 'time taken'; with the autistic sample taking significantly longer then the controls to complete the task. This is in line with previous research that has shown processing speed to be negatively affected in those with autism (Calhoun & Mayes, 2005; Oliveras-Rentas, Kenworthy, Roberson, Martin, & Wallace, 2012). However, given its poor psychometric properties, this task is an unlikely candidate for future development.

### *5.2.3 Alternating Sequence Task*

The Alternating Sequence task showed more promise with regards to its construct and criterion validity. Those with autism took longer to complete the task, made more additions to the patterns and made more overall errors than the control group. The 'total additions' and 'total errors' measures correlated significantly with the BRIEF GEC score whilst 'total additions' also correlated significantly with the BADS-C Zoo Map test (2). This would seem to suggest that the Alternating Sequence task, in particular the 'total additions' measure possesses criterion validity when compared with existing measures of executive function. Further, the measures both correlated with SCDC and SDQ scores which pertain to everyday autistic and psychological symptomology. Again, these findings are considered tentatively given multiple comparisons made in the study and the possibility of type 1 error. However, they suggest that the Alternating Sequence task may demonstrate veridicality and should therefore be considered for inclusion in any future task development.

As with the Pattern Drawing task, differences in 'time taken' to complete the task may simply reflect the slower processing speeds of those with autism. However, if we consider the demands of the Alternating Sequence task, it is possible to hypothesis which executive abilities may be influencing performance on the other measures. To perform the task correctly, the participants needed to attend to the researcher, remember how the sequence was drawn and replicate the pattern without perseveration. The results suggested that children with autism failed to do this as well as the typically developing controls, making more additions and using unusual strategies to recreate the pattern rather than following the researcher's demonstration. Written sequence tasks are known to draw on executive abilities associated with set

shifting including working memory, attention and inhibitory control (Arciniegas & Beresford, 2001). The autistic sample may have performed more poorly on this task because of impairment in these domains.

One limitation of this task was that the measure 'total additions' did not accurately capture whether these additions were new shapes or perseverations of the previous shapes. This information is important given that they may pertain to different areas of executive control (i.e. indicating impairments in inhibition versus working memory). Future development of the measure may include changes to the scoring to reflect this.

### *5.2.4 Storytelling Task*

When children with autism were asked to recount a story or to give an account of the testing session, they recalled a significantly fewer number of events than the controls. As such they took less time to recall the stories and made fewer errors in the sequence of events (given that they had less chance to make such errors). It would appear then that the primary outcome variable 'total number of events recalled', best captured differences between those with and without autism and provide evidence for the tasks construct validity. There were also significant differences for 'total number of confabulations' and 'total unclear characters' (number of times that a child referred to a character without it being clear who they were referring to). The Storytelling measures correlated with the BRIEF and SCDC scores and with the SDQ score (except for 'total confabulations'). There was also a significant correlation between 'total unclear characters' and the BADS-C Zoo Map (1) test. These findings provide some support for the criterion validity and veridicality of the Storytelling Task.

The ability to narrate a story or event involves linguistic, cognitive and socio-cognitive abilities (Tager-Flusberg & Sullivan, 1995). The finding that narrative length differed between the autistic and typically developing children contradicts previous studies which have failed to find differences in performance (e.g. Capps, Losh, & Thurber, 2000; Tager-Flusberg & Sullivan, 1995). Recalling fewer events might pertain to impairments in working memory as has been shown elsewhere (Wang et al., 2017). However, given that some studies have shown intact working memory capacities relating to articulatory rehearsal (Russell, Jarrold, & Henry, 1996), poorer performance on this task by the clinical group may also reflect difficulties with the organisation and sequencing of ideas and the pragmatic use of language (Diehl, Bennetto, & Young, 2006). For example, the children with autism may have failed to recognise the importance of key events in the story. The fact that the researcher had difficulty following autistic accounts (as indicated by 'total unclear characters') is also in line with previous findings (Colle, Baron-Cohen, Wheelwright, & Van der Lely, 2008). It may relate to deficits in 'theory of mind' something Russell views as a consequence of executive disorder (Russell, 1997). Finally, the fact that autistic children made more confabulations on the task is in line with expectations and may suggest impaired source memory and executive function (Spitzer, White, Mandy, & Burgess, 2017).

The Storytelling task was the most difficult task in the Eco-TED to standardise and devise a scoring system. As such there are several improvements that could be made to the task. Firstly, given that structural coherence is a key measure of narrative abilities (Baesler, 1995), it would make more sense to calculate 'percentage errors' rather than 'total errors' for sequence of events. This is because the controls recalled

significantly more events than the clinical sample creating a confounding effect. There is also evidence to suggest that children are more likely to give an in-depth narrative to someone that has not been present in the testing session (Liles, Duffy, Merritt, & Purcell, 1995). When we developed the task, it was difficult to see how this might be facilitated in practice. However, the fact that the child recalled the stories to the researcher may have influenced the results. One final point is that inter-rater reliabilities were not assessed in the current study. Given that all participant accounts for the Storytelling task were rated by one researcher, it would be important to assess whether the results obtained were influenced by rater bias. Therefore, the use of several rater's and the calculation of inter-rater reliabilities should be considered in the future.

### 5.3 Improving on Psychometric Properties

The Eco-TED may benefit from the addition of a questionnaire-based measure. This is something already included alongside tests such as the BADS-C, as they been shown to increase the sensitivity of measuring executive function deficits compared with novel problem solving tasks alone (Channon et al., 2001). Including such a measure in the Eco-TED might contribute to its veridicality.

The Eco-TED tasks were developed based on parental responses to an existing psychometric measure: the 3Di questionnaire. It could be argued that to take a truly bottom-up approach to development, tasks should have been based on those difficulties described by parents and caregivers first-hand. To do this we might have held focus groups or interviews to ascertain which behaviours were most troublesome. The advantage of this approach would be to develop tasks around the behaviours causing most functional impairment, which in turn might lead to more specific interventions.

## 5.4 Strengths and Limitations of the study

The Eco-TED measure was developed using a function-led approach, based on those impairments displayed by children with autism. As such it could be argued that one of the strengths of the Eco-TED is that is possesses greater verisimilitude than other measures currently available. This is because development of each of the Eco-TED tasks was based around the cognitive demands that exist for children with autism in their everyday lives. Comparison of the four tasks discussed in this study with existing measures would also suggest that the Eco-TED possesses a degree of veridicality, as demonstrated by correlations with existing tests of everyday symptom severity (e.g. the SDQ). Both of these criteria are essential to an ecologically valid measure (Kenworthy et al., 2008). Including a measure of adaptive functioning in any future assessments of the Eco-TED, such as the Vineland Adaptive Behaviour Scales (Sparrow, Cicchetti, & Balla, 2005), may contribute further evidence towards it's veridicality.

The fact that the tasks borne through a more iterative bottom-up process were better at discriminating between the autistic and control samples, promotes this method of task development. For example, the Luria test which has been used widely in neuropsychological assessment but was adapted for the Eco-TED, did not give rise to group differences. This would seem to lend support to the argument that current measures of executive function are not 'fit for purpose' when examining executive abilities in those with autism; and that more bespoke measures developed through a function led approach are required (Burgess et al., 2006).

There were clear limitations with the current study with the most pertinent relating to power. Creation and assessment of the validity of other ecologically valid

executive measures have had a sample size of 30-40 clinical and non-clinical participants (e.g. Norris & Tate, 2000). A sample size of 45 per group would have allowed sufficient power (.80) to detect a large effect size for test-correlations at the .01 level (Cohen, 1992). This was important given that we made multiple comparisons. Further, we adopted an exploratory approach to analysis as all data is considered to be informative in the initial stages of task development (Anastasi & Urbina, 1997). However, Nunnally (1978) suggests that studies should include twice as many participants to variables. A larger sample would have allowed us to address these concerns. Unfortunately, due to financial, time and recruitment constraints this was not possible. Future development of the measure would benefit from replication in a larger sample.

Previous criticisms of studies in this field have included concerns about inappropriate matching of samples, particularly concerning age and IQ (Kenworthy et al., 2008). Although the samples in our study were matched regarding these criteria, we did not control for co-morbid indications. Six of the participants had co-morbid disorders (including ADHD, ADD, OCD) that are known to have frontal lobe involvement or may have affected test performance (e.g. GAD) (Cobert, 2013). Therefore, impaired performance on the Eco-TED tasks may have been a consequence of their co-morbid presentations rather than being autism-specific. Our aim was to create a good measure that was ecologically valid rather than drawing strong conclusions about which executive constructs are involved. Therefore, for the purposes of our study, including children with co-morbid diagnoses was acceptable. Future development of the measure might want to consider including a clinical sample with only a primary diagnosis of autism and a sample of children with a co-morbid

diagnosis as well. This would indicate whether the Eco-TED tasks are measuring autism specific difficulties only.

Finally, there is the consideration of our samples and the generalisability of the findings. Firstly, we only included children that had a normal range of intellectual functioning. This was so that we could match our clinical and control samples adequately and to control for potential confounders such as verbal ability. However, it is widely known that there is a higher prevalence of intellectual disability amongst those with autism (Baird, 2006; Deb & Prasad, 1994; Fombonne, Quirke, & Hagen, 2011). The measure would need to be applied to those children with a FSIQ less than 70 to see whether differences in performance between groups are preserved. This is particularly pertinent to the Storytelling task which has a large verbal component. Secondly, the control sample was recruited predominantly from one school that had a relatively narrow demographic of white middle-class children. Seeking to include a wider demographic of children in the control sample would have been favourable, as would the inclusion of a higher number of girls. There were significantly more girls in the clinical group than the control group. Having a clinical group comprised of 50% females does not accurately reflect the true male-to-female ratio of autism, currently thought to be 3:1 (Loomes, Hull, & Mandy, *in press*). Future studies should address this given that the behavioural phenotype of autism is hypothesised to differ amongst males and females (Lai et al., 2011). This would have important implications for the measurement of any impairments.

# 6 Conclusion and Future Directions

This paper reports an exploratory study that's primary aim was to develop and pilot an ecologically valid measure of executive function. The tasks were specifically

designed for those with autism using a function-led approach. This paper reviewed the psychometric properties of four tasks. Two showed the potential to differentiate between those with and without autism but would benefit from replication in a larger sample to confirm findings and further assess construct validity. When compared with measures of symptom severity and existing measures of executive functioning, findings were mixed. This suggests a need to improve the validity and reliability of the Eco-TED but may also reflect the limitations of existing measures. Developing a measure using this function-led approach is an intensely iterative process that is not without its limitations. However, the creation of ecologically valid measures that resemble everyday cognitive demands and are thus better at predicting functioning, have obvious benefits for future assessment and intervention.

# 7 References


Abdi, H. H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. *Encyclopedia of Measurement and Statistics.*, *1*, 1–9.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*(1), 17–42.

American Psychiatric Association. (2013). *DSM-V*. *Diagnostic and Statistical Manual of Mental Disorders 5th Edition*. Washington, DC,: Author.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). New Jersey: Prentice Hall.

Anderson, V. (1998). Assessing Executive Functions in Children: Biological, Psychological, and Developmental Considerations. *Neuropsychological Rehabilitation*, *8*(3), 319–349.

Anglim, J. (2009). No Title. Retrieved February 20, 2017, from http://jeromyanglim.blogspot.co.uk/

Arciniegas, D. B., & Beresford, T. P. (2001). *Neuropsychiatry: An Introductory Approach*. Cambridge: Cambridge University Press.

Baesler, E. J. (1995). Construction and test of an empirical measure for narrative coherence and fidelity. *Communication Reports*, *8*(2), 97–101.

Baird, G. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215.

Barker, C., Pistrang, N., & Elliott, R. (2002). *Research methods in clinical psychology: An introduction for students and practitioners*. *Methods*.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive*

*Sciences*, *6*(6), 248–254. http://doi.org/10.1016/S1364-6613(02)01904-6

Baron-cohen, S., Leslie, A. M., & Frith, U. T. A. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Baron, I. S. (2007). Test Review: Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C). *Child Neuropsychology*, *13*(6), 539–542.

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. 2: Response inhibition. *Autism : The International Journal of Research and Practice*, *9*(1), 29–43.

Bristow, E. (2016). *Developing and Piloting a New Measure of Executive Functioning for Children with Autism Spectrum Disorder (ASD)*. University College London.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., … Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society : JINS*, *12*(2), 194–209.

Calhoun, S. L., & Mayes, S. D. (2005). Processing speed in children with clinical disorders. *Psychology in the Schools*, *42*(4), 333–343.

Capps, L., Losh, M., & Thurber, C. (2000). No Title. *Journal of Abnormal Child Psychology*, *28*(2), 193–204.

Channon, S., Charman, T., Heap, J., Crawford, S., & Rios, P. (2001). Real-Life-Type Problem-Solving in Asperger ' s Syndrome, *31*(5).

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills.

*Neuropsychology Review*, *13*(4), 181–97.

Cobert, J. (2013). *Tarascon Pediatric Psychiatrica*. Burlington, MA: Jones and Bartlett Learning.

Cohen, J. (1992). Cohen 1992 Power Primer.pdf. *Psychological Bulletin*, *112*(1), 155–159.

Colle, L., Baron-Cohen, S., Wheelwright, S., & Van der Lely, H. K. J. (2008). Narrative discourse in adults with high-functioning autism or Asperger syndrome. *Journal of Autism and Developmental Disorders*, *38*(1), 28–40.

Crawford, J. R., & Garthwaite, P. H. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: evaluation of alternatives using monte carlo simulations and revised tests for dissociations. *Neuropsychology*, *19*(3), 318–31.

Davis, A. S. [Ed]. (2011). Handbook of pediatric neuropsychology. *Handbook of Pediatric Neuropsychology.* New York, NY, US: Springer Publishing Co.

Deb, S., & Prasad, K. B. (1994). The prevalence of autistic disorder among children with a learning disability. *The British Journal of Psychiatry*, *165*(3), 395–399.

Diamond, A. (2001). A Model System for Studying the Role of Dopamine in the Prefrontal Cortex during Early Development in Humans: Early and Continuously Treated Phenylketonuria. In C. A. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 433–472). Cambridge, MA, US: A Bradford Book.

Diehl, J. J., Bennetto, L., & Young, E. C. (2006). Story Recall and Narrative Coherence of High-Functioning Children with Autism Spectrum Disorders, *34*(1), 87–102.

Elliott, R. (2003). Executive functions and their disorders. *British Medical Bulletin*, *65*, 49–59.

Emslie, H., Wilson, C. F., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2003). Behavioural Assessment of the Dysexecutive Syndrom in Children (BADS-C). Pearson Clinical.

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.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). London, England: SAGE Publications Inc.

Fombonne, E., Quirke, S., & Hagen, A. (2011). Epidemiology of pervasive developmental disorders. In D. . Amaral, G. Dawson, & D. H. Geschwind (Eds.), *Autism Spectrum Disorders* (pp. 90–111). New York: Oxford University Press.

Franzen, Michael D., and K. L. W. (1996). *Conceptual Foundation of Ecological Validity in Neuropsychological Assessment.*

Frith, U. (2003). *Autism: explaining the enigma* (2nd ed.). Malden, MA: Blackwell.

Geurts, H. M., Corbett, B., & Solomon, M. (2009). The paradox of cognitive flexibility in autism. *Trends in Cognitive Sciences*, *13*(2), 74–82.

Gidley Larson, J. C., & Mostofsky, S. H. (2008). Evidence that the pattern of visuomotor sequence learning is altered in children with autism. *Autism Research*, *1*, 341–353.

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*, 235–238.

Gioia, G. a, & Isquith, P. K. (2004). Ecological assessment of executive function in traumatic brain injury. *Developmental Neuropsychology*, *25*(1–2), 135–58.

Glass, G. ., Peckham, P. ., & Sanders, J. . (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effect Analyses of Variance and Covariance. *Review of Educational Research*, *42*(3), 237–288.

Golden, C. J. (1991). Luria-Nebraska neuropsychological battery: Children' s revision manual. Los Angeles: Western Psychological Services.

Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive Functions in Young Children with Autism. *Child Development*, *70*(4), 817–832.

Happe, F. (1999). Autism: cognitive deficit or cogntive style? *Trends in Cogntive Science*, *3*(6), 216–222.

Henry, L. a., & Bettenay, C. (2010). The Assessment of Executive Functioning in Children. *Child and Adolescent Mental Health*, *15*(2), 110–119.

Henry, L. A., & 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. http://doi.org/10.1016/j.dr.2004.01.001

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), 2822–35.

Homack, S. R., & Reynolds, C. R. (2007). *Essentials of assessment with brief intelligence tests*. *Psychological Assessment*.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction

in autism. *Neuropsychologia*, *32*(4), 477–492.

Irby, S. M., & Floyd, R. G. (2013). Test Review: Wechsler Abbreviated Scale of Intelligence, Second Edition. *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), 320–38.

Kvavilashvili, L., & Ellis, J. (2004). Ecological validity and the real-life/laboratory controversy in memory research: A critical (and historical) review. *History and Philosophy of Psychology*, *6*, 59–80.

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).

Leark, R. A. (2004). The Luria-Nebraska Neuropsychological Battery – Children ' s Revision. In G. Goldstein & S. Beers (Eds.), *Comprehensive Handbook of Pschological Assessment, Volume 1: Intellectual and Neuropsychological Assessment.* (Vol. 1). John Wiley & Sons Inc.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.)*. *Neuropsychological assessment (5th ed.).* Oxford: Oxford University Press.

Liles, B. Z., Duffy, R. J., Merritt, D. D., & Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research*, *38*(April 1995), 415–425.

Liss, M., Fein, D., Allen, D., Dunn, M., Feinstein, C., Morris, R., … Rapin, I. (2001).

Executive functioning in high-functioning children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 261–70.

Loomes, R., Hull, L., & Mandy, W. (n.d.). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*.

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.

Luria, A. R. (1980). Higher Cortical Functions in Man. New York, NY: Basic Books.

McCrimmon, A. W., & Smith, A. D. (2012). Review of the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). *Journal of Psychoeducational Assessment*, *31*(3), 337–341.

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 Exp Neuropsychol*, *14*(5), 749–761.

Miziara, C., Giraldes de Manreza, M., Mansur, L., Conti Reed, U., & Buchpiguel, C. (2013). Sequential motor task (Luria's Fist-Edge-Palm Test) in children with benign focal epilepsy of childhood with centrotemporal spikes. *Arquivos de Neuro-Psiquiatria*, *71*(6).

Moses Jr., J. A., Golden, C. J., Berger, P. A., & Wisniewski, A. M. (1981). Neuropsychological deficits in early, middle, and late stage Huntington's disease as measured by the Luria-Nebraska Neuropsychological Battery. *Int.J.Neurosci.*,

*14*(1–2), 95–100.

Mostofsky, S. H., Goldberg, M. C., Landa, R., & Denckla, M. B. (2000). Evidence for a deficit in procedural learning in children and adolescents with autism: Implications for cerebellar contribution. *Journal of the International Neuropsychological Society*, *6*(7), 752–759.

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.

Nunnally, J. C. (1978). *Psychometric Thoery* (2nd ed.). New York: McGraw-Hill.

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. http://doi.org/10.1007/s10803-011-1289-7

Owen, W. W. (1985). Creating a Summary Table of Normalised (Z) Scores. Retrieved February 20, 2017, from http://www.sascommunity.org/sugi/SUGI85/Sugi-10-199 Owen.pdf

Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, *9*(4), 491–500.

Ozonoff, S., Pennington, Bruce, F., & Rogers, Sally, J. (1991). Executive function deficits in high-functioning autistic.pdf. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *32*(7), 1081–1105.

Pastor-Cerezuela, G., Fernandez-Andres, M.-I., Feo-Alvarez, M., & Gonzalez-Sala, F.

(2016). Semantic Verbal Fluency in Children with and without Autism Spectrum Disorder: Relationship with Chronological Age and IQ. *Frontiers in Psychology*, *7*, 921.

Pennington, Bruce, F., Rogers, Sally, J., Bennetto, L., McMahon Griffith, E., Reed, T., & Shyu, V. (1997). Validity tests of the executive dysfunction hypothesis of autism. In J. Russell (Ed.), *Autism as an executive disorder.1* (pp. 143–178). Oxford: Oxford University Press.

Pennington, B. F., & Ozonoff, S. (1996). Executive Functions and Developmental Psychopathology. *Journal of Child Psychology and Psychiatry*, *37*(1), 51–87.

Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, *27*(2), 224–260.

Robbins, T. W. (1997). Integrating the neurobiological and neuriphyssiological dimensions of autism. In J. Russell (Ed.), *Autism as an executive disorder.* (pp. 21–56). Oxford: Oxford University Press.

Robinson, S., Goddard, L., Dritschel, B., Wisley, M., & Howlin, P. (2009). Executive functions in children with autism spectrum disorders. *Brain and Cognition*, *71*(3), 362–8.

Roy, A., Allain, P., Roulin, J. L., Fournet, N., & Le Gall, D. (2015). Ecological approach of executive functions using the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C): Developmental and validity study. *Journal of Clinical and Experimental Neuropsychology*, *37*(9), 956–971.

Roy, E., Bottos, S., Pryde, K., & Dewey, D. (2004). Approaches to understanding the neurobehavioural mechanisms associated with motor impairments in children. In D. Dewey & D. E. Tupper (Eds.), *Developmental Motor Disorders: A*

*Neuropsychological Perspective* (pp. 44–65). New York, NY: Guilford Press.

Rumsey, J. M., & Hamburger, S. D. (1988). Neuropsychological findings in high-functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology*, *10*(2), 201–221.

Russell, J. (1997). *Autism as an executive disorder*. *New York NY US Oxford University Press* (Vol. 313).

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *37*(6), 673–686.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727–741.

Shallice, T., & Evans, M. E. (1978). The involvement of the frontal lobes in cognitive estimation. *Cortex*, *14*, 294+303.

Skuse, D. H. (2013). No Title, (2005), 2007–2008.

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*(June), 705–708.

Skuse, D. H., Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits : heritability , reliability and validity of the Social and Communication Disorders Checklist. *British Journal of Psychiatry*, *187*, 568–573.

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), 548–58.

Sparrow, S. S., Cicchetti, D. V, & Balla, D. A. (2005). Vineland Adaptive Behavior Scales, Second Edition (Vineland-II). Pearson Clinical.

Spitzer, D., White, S. J., Mandy, W., & Burgess, P. W. (2017). Confabulation in children with autism. *Cortex*, *87*, 80–95.

Tager-Flusberg, H., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics*, *16*, 241.

Teunisse, J.-P., Roelofs, R. L., Verhoeven, E. W. M., Cuppen, L., Mol, J., & Berger, H. J. C. (2012). Flexibility in children with autism spectrum disorders (ASD): inconsistency between neuropsychological tests and parent-based rating scales. *Journal of Clinical and Experimental Neuropsychology*, *34*(7), 714–23.

Towgood, K. J., Meuwese, J. D. I., Gilbert, S. J., Turner, M. S., & Burgess, P. W. (2009). Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. *Neuropsychologia*, *47*(13), 2981–8.

Tukey, J. W. (1960). A Survey of Sampling from Contaminated Distributions. *Contributions to Probability and Statistics Essays in Honor of Harold Hotelling*, *2*, 448–485.

Umetsu, A., Okuda, J., Fujii, T., Tsukiura, T., Nagasaka, T., Yanagawa, I., … Yamadori, A. (2002). Neural correlates of auditory perception in Williams syndrome: an fMRI study. *NeuroImage*, *17*, 385–392.

Wang, Y., Zhang, Y., Liu, L., Cui, J., Wang, J., Shum, D., … Chan, R. (2017). A Meta-Analysis of Working Memory Impairments in Autistic Spectrum Disorders. *Neuropsychological Review*.

Weiner, M. F., Hynan, L. S., Rossetti, H., & Falkowski, J. (2011). Luria's three-step

test: what is it and what does it tell us? *International Psychogeriatrics*, *23*, 1602–1606.

Weschler, D. (1999). WASI: Wechsler abbreviated scale of intelligence. Hove: Psychological Corporation.

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research : Official Journal of the International Society for Autism Research*, *2*(3), 138–47.

Wilkinson, L. A. (2010). *A best practice guide to assessment and intervention for autism and Asperger syndrome in schools*. London: Jessica Kingsley Publishers.

Wilson, B. a., Evans, J. J., Emslie, H., Alderman, N., & Burgess, P. (1998). The Development of an Ecologically Valid Test for Assessing Patients with a Dysexecutive Syndrome. *Neuropsychological Rehabilitation*, *8*(3), 213–228.

Wood, R. L., & Liossi, C. (2006). The ecological validity of executive tests in a severely brain injured sample. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *21*(5), 429–37.

Youth in Mind. (n.d.). Strengths and Difficulties Questionnaire. Retrieved February 1, 2017, from http://www.sdqinfo.org

# Part 3: Critical Appraisal

# 1. Introduction

Some have likened the research process to a 'journey' characterised by periods of optimism and excitement but also marred by moments of risk and disappointment (Barker, Pistrang, & Elliott, 2002; Kvale, 1996). In the end, it is hoped that this process leads to a piece of work that develops our understanding and contributes towards our scientific knowledge. This appraisal is a reflection on the research process I undertook including; background to the project, key learning points from conducting discovery-orientated research; and future directions of the Eco-TED. It also considers my research and experiences in the context of an evidence-based approach. Finally, it closes with a discussion on what I have learnt about Autistic Spectrum Disorder (ASD).

# 2. Origins of the Research Project

I embarked on this research project with very little experience of working with children with ASD. Most of my prior knowledge about the disorder stemmed from my work with Adults for whom ASD was a co-morbid diagnosis alongside intellectual disability. However, what this work had led me to understand was the true pervasiveness of the condition. Many individuals with ASD require continued care for the duration of their lives, significantly impacting families and having important implications for care providers (Karst & Vaughan Van Hecke, 2012; Seltzer, Greenberg, Floyd, Pettee, & Hong, 2001; Volkmar & Pauls, 2003). During my time as an assistant psychologist I had the opportunity to attend a forum led by parents of children with autism. I was fascinated by the experiences of those families, some of whom had children that were not able to communicate verbally whilst others had

children that were comparatively high functioning. What was clear was that all of the accounts shared a common experience of parental stress as a consequence of their child's symptoms, something which has been reported in the literature (Rao & Beidel, 2009).

I saw the current research project as an opportunity to contribute towards a rapidly expanding field of clinical psychology. I was aware of the theoretical models of autism but I wasn't sure how useful these models were in clinical practice. This became especially apparent when on placement in a child neuropsychological setting. It was possible to assess and diagnose ASD using a range of pre-existing measures. However, they did not always translate well to parental report of the child's functioning and therefore were not always helpful in designing interventions. So, by getting involved in a project whose aim was to develop a measure of ASD that had better ecological validity, I was hoping to contribute towards both the theoretical understanding of ASD and clinical practice. Ultimately, like most researchers set out to do, my intention was to produce a meaningful piece of research that might make some difference to those children and families affected by ASD.

# 3. The Research Process

## 3.1 The Research Team

Along with a service related assignment, this research project was my first experience of research from inception through to dissemination. Although I had previous experience of research, this had been gained on large-scale, multi-centre projects that were already well established. As such, the prospect of creating a solid research proposal, gaining ethics, forming relationships with services from which to recruit participants and carrying out the research seemed a daunting task. However,

being part of a research team which comprised two experienced research supervisors and another doctoral student, allayed some of these fears.

One of the key learning points from engaging in team research is the usefulness of having others to consult, resolve problems and develop ideas with (Barker et al., 2002; Hodgson & Rollnick, 1996). This was particularly true for the current research project given its bottom-up approach to task development which essentially meant 'starting from scratch'. The different clinical and research experiences of those in the team meant that we could approach the design and development of the Eco-TED from both an academic and clinical stance. For example, whilst brainstorming tasks that related to the 3Di items, I was encouraged to hold in mind the scoring of my proposed tasks and to pursue only those tasks that might afford better clinical utility. At the same time, I could contribute knowledge and experience gained through working in a paediatric neuropsychology setting, such as how to engage children in the testing process and which tasks children might be more likely to enjoy. Taking a collaborative approach of this nature meant that the research process remained more focused than it may have otherwise. It also meant that during periods of being overwhelmed by confusing alternatives, I could seek the counsel of my colleagues (Kvale, 1996).

### 3.2 Ethics and Permissions

As the project utilised a clinical group recruited through a specialist NHS service, both ethics and site specific permissions had to be sought. This was arguably the most time-consuming part of the whole process. However, it also encouraged deeper thinking about the protection of those taking part in studies.

Research fatigue is something that has been highlighted elsewhere, particularly amongst those groups that are more heavily researched (Clark, 2008, 2010). Although

all the families that we approached had consented to be contacted for research purposes, it was important that we respected the demands that taking part in the study might place on both the children and their families. Although small incentives were offered, it was clear from speaking to the families that their participation was borne out of perceived future benefit to themselves or others, interest in the study outcomes and altruism. These are all reasons that are known to motivate individuals to take part in research (Clark, 2010). However, research has also shown that failing to keep participants informed of outcomes or acknowledge their contribution can negatively impact their impressions of taking part in studies (Barker et al., 2002). For reasons beyond our control dissemination of research findings took longer than anticipated. It would be interesting to know whether families had an overall positive experience of taking part in our study despite this, particularly as future development of the measure would most likely rely on recruitment from the same participant pool.

For confidentiality reasons, clinical team members that were known to the families made the initial research contact. Given that these professionals were the first contact that families had with the research project it was essential that they had sufficient knowledge of the study but also that they were supportive of the study's aims. As the 'gatekeepers' within the service, these professionals were instrumental in ensuring access and successful recruitment of participants (Benton & Cormack, 2000). Given the ever-increasing demands on NHS services and resources, it was important that we balanced promotion of our research project within the service whilst limiting the demands placed upon clinicians. Regular liaison and attendance at team meetings was time-consuming but something that was necessary in developing successful working

relationships between research groups and NHS settings (Swan, Robertson, & Evans, 2009).

The above is also true for the recruitment of our control sample, most of which we recruited from local schools. Unlike the NHS setting, head-teachers had less immediate benefit from our research. Therefore, it was necessary to help them realize what could be gained from allowing access to their students. For example, some asked for psychoeducational information relating to ASD or other aspects of psychological wellbeing. I underestimated the importance of this initial groundwork and networking. In the future, I would consider aligning the goals of the research with the goals of the setting more thoroughly in the planning stages (Hardy, 1993) to ensure faster and more efficient recruitment.

### 3.3 Continuity of Data Collection

Due to unforeseen circumstances, it was necessary to enlist the help of three Research Assistants in the data collection stage of the project. These research assistants were all trained in the administration of the Eco-TED and the additional neuropsychological measures. This training was undertaken by myself and a fellow doctoral student and involved each research assistant observing two test administrations and being observed administering the tests themselves. Although attempts were made to standardize test administration in this way, there is a question of how this might have affected the results of the study.

Unlike pre-existing neuropsychological measures, the Eco-TED was being administered for the first time. Rigorous standardisation is essential to maintaining the validity of any neuropsychological test (Strauss, Sherman, & Spreen, 2006). It could be argued that using just one or two administrators at this stage would have better

managed the issue of standardisation. This is particularly true given that small changes in task format or administration can lead to very different demands being placed on executive abilities (Stuss et al., 2000; Stuss, Binns, Murphy, & Alexander, 2002). Further, it has been suggested that the presence of a third-party during test administration, even as an observer, can lead to social facilitation effects which impact neuropsychological test results (Shindell, McCaffrey, & Silk-Eglit, 2014). Therefore, by observing the research assistants administer the battery for some participants, it was possible that we introduced a social confounder. This might have been particularly pertinent for the clinical sample given that performance of children with autism is known to be affected by socio-cognitive demands (Kenworthy, Yerys, Anthony, & Wallace, 2008).

# 4 Developing a neuropsychological measure

## 4.1 Adopting a function-led approach

Our aim was to produce an ecologically valid measure that possessed both verisimilitude and veridicality (Kenworthy et al., 2008). To do this we used a function-led or bottom-up approach to designing the Eco-TED, whereby we based tasks on those impairments already reported in children with ASD. In doing so we hoped that our tasks would resemble the cognitive demands that the child would face in everyday life and that consequently performance on the Eco-TED would correlate more highly with everyday functioning.

Initially this approach was very appealing because we were working at the level of directly observable behaviour rather than having to think about complex theoretical underpinnings (Burgess et al., 2006). So, the task was simply to design a measure based on this observed behaviour (using 3Di data). However, it was a method

that was unfamiliar to me as much of my previous teaching had emphasised a theoretical, construct-driven approach. Borrowing from the phenomenology literature, I found it hard at times to 'bracket' my pre-existing knowledge of proposed executive function deficits in ASD and the neuropsychological tests already in existence to measure these. Although the tests that we included in the final battery were all borne from a bottom-up process, two of my tasks were developed from measures previously used in the neuropsychological literature (Luria and Alternating Sequence Task). The use of pre-existing tasks as a starting point in the development of ecologically valid measures is evidenced elsewhere (e.g. Wilson, Evans, Emslie, Alderman, & Burgess, 1998). Yet, arguably this emphasises the difficulty in creating a truly novel measure that does not draw on tests originally developed for a different population.

The bottom-up approach also highlighted some important points concerning the difficulty of integrating clinical utility and theoretical understanding. Throughout designing the measure, we hypothesised which aspects of executive function might be implicated in the tasks that we were designing and the everyday activities they pertained to. Although developing an ecologically valid measure was the main aim of our project, we also wanted it to contribute towards scientific understanding of executive dysfunction in ASD. We made tentative links between our tasks and possible executive impairments based on our theoretical knowledge. For example, we posited that working memory impairments, socio-cognitive deficits and an inability to integrate information may be linked to poorer performance on the storytelling task. However, it is not possible to draw stronger conclusions about the exact executive constructs that might be involved. One way that researchers have developed this construct level understanding for ecologically valid tests in the past is to link

performance of the tasks to brain structure as an additional step in development (Burgess et al., 2006). Building up levels of explanation in this way might be one future direction of the Eco-TED; increasing its contribution to our understanding of executive dysfunction.

### 4.2 Dilemmas of ecological validity

There is good evidence that existing measures of executive function do not sufficiently capture the difficulties seen in everyday life in those with autism (Burgess et al., 2006; Channon, Charman, Heap, Crawford, & Rios, 2001; Hill & Bird, 2006). Some say that this is because traditional 'lab-based' tests assess epiphenomenon's that do not necessarily express themselves in the real world (Burgess et al., 2006). Further, real-life impairments may only be seen when a combination of factors vary together in specific ways (Kingstone, Smilek, Birmingham, Cameron, & Bischof, 2005). We argued that developing an ecologically valid measure from a bottom-up approach would allay some of these problems. However, in doing so it also raised some important points.

Firstly, what was clear when designing the Eco-TED was how difficult it is to maintain ecological validity of tasks whilst incorporating them in to an assessment battery that was administered in a formalised way. Some early ideas for task prototypes seemed to be closer to the original 3Di items than later ones that were included in the battery. An example is the 3Di item "Is X able to tie his/her shoes without help?". We developed some prototypes early on which used an actual shoe as the stimulus and required the child to make specific patterns with the laces. When this task was piloted it became clear that it would be difficult and time consuming to administer. Despite this task feeling more ecologically valid than some of the other tasks (e.g. the Luria

task), it was not practical to include it in the battery. This raises the question of whether it is possible to create a truly ecologically valid measure that is quick and easy to administer, yet maintains the complex interplay of factors that would be present in real-life.

A second question is whether it is possible for a measure to be ecologically valid when it is administered by a researcher or clinician in a controlled testing environment? This setting is far from a naturalistic environment and may reduce cognitive demands through excessive structure and cues (Silver, 2000). One of the criticisms of more modern tests of executive function which do possess ecological validity is that they lack the standardisation of more traditional neuropsychological tests. For example they have been shown to demonstrate poor reliability and validity (Henry & Bettenay, 2010). Indeed test-retest reliability for our measure was far below that which would be expected. Therefore, there is the dilemma as to whether it is possible to maintain scientific rigour and standardisation whilst also creating something which is novel, lifelike and clinically useful.

### 4.3 Piloting

One way to enhance scientific rigour is through appropriate piloting. This is seen as an essential stage of any task or intervention development (Feeley et al., 2009). Although our project did include some informal piloting of our measure, the piloting sample consisted solely of typically developing children. The piloting process enabled us to refine tasks for the final battery. However, it was clear when we administered the battery on the clinical sample that there were some things that we had not considered. For example, in general the testing session took longer for clinical participants. The children with autism required more breaks, more time spent building rapport and they

were more likely to refuse tasks if they became fatigued or found a task difficult. Particularly in the case of the latter, piloting on children with an ASD diagnosis would have allowed us to plan administration more carefully. For example, adding discontinuation rules into the task script in the event of task refusal and allowing more frequent breaks from the outset to prevent testing fatigue. It may also have helped us to identify those tasks that were likely to show a ceiling effect (i.e. the Luria task) so that we could exclude them from the battery or adapt them accordingly.

When piloting on typically developing children, the verbal feedback was generally positive with children describing the tasks as "fun". However, one thing we did not do was to ask our study participants to rate their experiences of completing the Eco-TED. This would have given us information about how the tasks were experienced, especially by those with autism. We considered an important part of ecological validity to be a measure that was enjoyable to complete; one which did not feel like an 'assessment'. Participant feedback on which of our tasks achieved this would have been useful information for the future development of the measure.

### 4.4 Task Development

This was my first experience of developing a task based neurospychological measure. I was lucky to have the support of two supervisors who had significant prior knowledge and experience of this. However, at times the process was overhwelming and frustrating. In particular, a large amount of thought and resources go in to the initial stages of task design. I spent significant time developing several task prototypes that were not taken forward for various reasons. During this process I realised how easy it is to become overly invested in a task. There is evidence of a confirmatory bias in experiemental studies, particularly concering hypothesis-determined information

seeking (Nickerson, 1998). Although the current study differed in that it was exploratory in nature, it still highlighted how easy it is to become wedded to particular ideas and hypotheses. In this instance having the support of a research team that could help me think more reflexively about the tasks was important.

The second issue that I encountered with regards to task development, was how to make the measures as ecologically valid as possible without becoming overwhelmed by outcome variables. A case in point is the storytelling task. Initially we thought that it might be possible to capture four of the 3Di items in this measure; (1) Would X have difficulty in explaining to a younger child how to play a simple game; (2) What about difficulty in telling a story or describing what (s)he has done; (3) Has X ever played a game with life-like figures or animals in which she talks to them and; (4) Has X ever played a game in which there are several figures or animals and they are talking to each other? The key to developing this task was to ensure that I was not simply re-producing existing semi-structured observational batteries such as the Autism Diagnostic Observation Scale (ADOS) (Lord et al., 1989) which involves observing children during play. It became clear that trying to capture all these items in one task was too difficult and led us to tasks that resembled the ADOS too closely. Therefore, I chose to focus the storytelling task on just items (1) and (2). Despite this, it was clear that the nature of the executive functions that this task drew on, namely narrative coherence and working memory, could not be captured by a simple scoring system. I was therefore posed with the dilemma of collecting enough information to make the task valuable but not so much information as to make the task unworkable. This taught me a lesson about the value of keeping tasks simple and parsimonious. It also made

me curious as to whether it is possible to create an ecologically valid measure that possessed these qualities.

## 5 Future Directions of the Eco-TED

Some of the tasks that we designed for the Eco-TED warrant further investigation. The Alternating Sequence task and the Storytelling tasks showed some ability to differentiate between the clinical and control group. However, further development and refinement of these measures are needed. They also need to be administered to a larger participant pool to increase the power of the findings (see Empirical Paper: Discussion). The initial aims of this project were to develop a measure that has greater verisimilitude and veridicality. This is something that has not yet been achieved if we consider the reliabilities and validities of the current measure. Therefore, future development may also want to consider how these can be improved upon, for example by using multiple data sources to determine executive impairments (Silver, 2000). Finally, the Eco-TED tasks were developed, analysed and written up as two separate projects but in the future it would be wise to consider the development and the validation of the Eco-TED battery as a whole (Russell, Russell, & Hill, 2005).

## 6 What I have learnt about Autistic Spectrum Disorder

ASD is a complex disorder that shows marked heterogeneity amongst those with a diagnosis. This is something I had limited knowledge of from my theoretical and clinical experience of ASD prior to commencing this project. However, the systematic review and development of the Eco-TED have both served to emphasise this point. For those Eco-TED tasks that indicated a group effect, those with ASD generally performed poorly compared with controls. However, it was also clear that

some of those with ASD performed similarly or better than their non-clinical peers. This may have been a consequence of the poor discriminant validity of our tasks (see below). Or it may replicate previous research findings that have demonstrated varying executive capabilities in those with the disorder. Regardless, an idiographic small-N approach that seeks to explore within-subject variance is one that is suited to this population as it prevents group effects from masking important individual differences (Shallice, 1979).

The above also highlights the difficulties inherent in designing executive measures specifically for those with ASD. Given that the population shows marked heterogeneity, it is difficult to create tasks that differentiate all those with ASD from typically developing peers. It is also true that executive impairments exist in those with a range of other neurodevelopmental and psychiatric disorders. This problem of discriminant validity is well recognised amongst the literature (see Ozonoff, 1997). Even for those tasks of the Eco-TED that did show some ability to discriminate between those with and without ASD, it is impossible to say whether this difference can be explained by ASD alone; or whether there are other factors and co-morbid reasons for impaired performance.

Something that was particularly striking when conducting this research was the dominant narrative of 'impairment' and 'dysfunction' surrounding those with ASD. As previously discussed, some studies have shown those with ASD to perform better than their typically developing counterparts (e.g. Lind & Bowler, 2010; Towgood, Meuwese, Gilbert, Turner, & Burgess, 2009). This raises the question of how helpful it is to pursue evidence of executive dysfunction in ASD and contribute towards this dominant narrative. Russell (1997) argues that understanding this phenomenon may

lead to better treatments and interventions as is the case in other disorders involving executive deficits, such as schizophrenia. Embarking on this project has highlighted to me the importance of maintaining a scientist-practitioner stance. Research is useless if it does not have practical application and go some way to improving the lives of those that it seeks to understand. Therefore, it is not enough to find evidence of executive impairment in children with ASD, even with ecologically valid tests. The key is translating this into suitable interventions and meaningful outcomes.

# 7 Conclusion

The complexity of ASD and the resulting difficulties in assessing executive function in those with ASD has been reinforced by this research. I have become conscious of the tendency for the literature to create a dominant discourse of impairment in those with ASD. Yet the degree of heterogeneity means that we might be missing some supra-normal abilities in this population. Despite this, those with ASD do possess difficulties in some aspects of their everyday lives associated with executive function. A function-led approach offers a chance to develop more ecologically valid measures to capture these. However, this is not an easy feat and the process raises important questions regarding the validity, reliability and standardisation of such measures.

Both the Alternating Sequence and Storytelling tasks warrant further investigation and possible development. This might be in the form of task improvement, changes to scoring and larger scale replication. The Eco-TED would also benefit from the development of a questionnaire based measure, particularly as this might help to increase its reliability and validity in practice.

This project has given me invaluable experience of conducting discovery-oriented research as part of an experienced research team. Reflecting on my experiences of this process has emphasised the circular nature of scientific research (Barker et al., 2002). Aside from the limitations I have discussed, this project has given rise to key questions and directions for future research and in doing so "adds voice to those who have gone before" (Barker et al., 2002).

# 8 References


Barker, C., Pistrang, N., & Elliott, R. (2002). *Research methods in clinical psychology: An introduction for students and practitioners*. *Methods*.

Benton, D. C., & Cormack, D. F. S. (2000). Reviewing and evaluating the literature. In D. F. S. Cormack (Ed.), *The research process in nursing* (4th ed.). Oxford: Blackwell Sciences Ltd.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M., Dawson, D. R., … Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society : JINS*, *12*(2), 194–209.

Channon, S., Charman, T., Heap, J., Crawford, S., & Rios, P. (2001). Real-Life-Type Problem-Solving in Asperger ' s Syndrome, *31*(5).

Clark, T. (2008). "We"re Over-Researched Here!': Exploring Accounts of Research Fatigue within Qualitative Research Engagements. *Sociology*, *42*(5), 953–970.

Clark, T. (2010). On "being researched": why do people engage with qualitative research? *Qualitative Research*, *10*(4), 423–438.

Feeley, N., Cossette, S., Cŏté, J., Héon, M., Stremler, R., Martorella, G., & Purden,

M. (2009). The importance of piloting an RCT intervention. *Canadian Journal of Nursing Research*, *41*(2), 84–99.

Hardy, G. E. (1993). Organisational issues: Making research happen. In M. Aveline & D. A. Shapiro (Eds.), *Research foundations for psychotherapy practice*. Chichester: Wiley and Sons.

Henry, L. A., & Bettenay, C. (2010). The Assessment of Executive Functioning in Children. *Child and Adolescent Mental Health, 15*(2), 110–119.

Hill, E. L., & Bird, C. M. (2006). Executive processes in Asperger syndrome: patterns of performance in a multiple case series. *Neuropsychologia*, *44*(14), 2822–35.

Hodgson, R., & Rollnick, S. (1996). More fun, less stress: How to survive in research. In G. Parry & F. N. Watts (Eds.), *Behavioural and Mental Health Research: A handbook of skills and methods* (2nd ed.). Hove: Lawrence Erlbaum Associates Publishers.

Karst, J. S., & Vaughan Van Hecke, A. (2012). Parent and Family Impact of Autism Spectrum Disorders: A Review and Proposed Model for Intervention Evaluation. *Clinical Child and Family Psychology Review*, *15*(3), 247–277.

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), 320–38.

Kingstone, A., Smilek, D., Birmingham, E., Cameron, D., & Bischof, W. F. (2005). Cognitive ethology: Giving real life to attention research. In J. Duncan, L. Phillips, & P. McLeod (Eds.), *Measuring the mind: speed, control and age* (pp. 341–357). Oxford: Oxford University Press.

Kvale, S. (1996). An introduction to qualitative research interviewing. *Sage*

*Publications*, 344.

Lind, S. E., & Bowler, D. M. (2010). Episodic memory and episodic future thinking in adults with autism. *Journal of Abnormal Psychology*, *119*(4), 896–905.

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.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), *Autism as an executive disorder.* (pp. 179–211). Oxford: Oxford University Press.

Rao, P. A., & Beidel, D. C. (2009). The Impact of Children with High-Functioning Autism on Parental Stress, Sibling Adjustment, and Family Functioning. *Behaviour Modification*, *33*(4), 437–451.

Russell, E., Russell, S., & Hill, B. (2005). The fundamental psychometric status of neuropsychological batteries. *Archives of Clinical Neuropsychology*, *20*(6), 785–794.

Russell, J. (1997). *Autism as an executive disorder*. *New York NY US Oxford University Press* (Vol. 313).

Seltzer, M., Greenberg, J. ., Floyd, F. J., Pettee, Y., & Hong, J. (2001). Life course impacts of parenting a child with a disability. *American Journal of Mental Retardation*, *106*(3), 265–286.

Shallice, T. (1979). Case study approach in neuropsychological research. *Journal of*

*Clinical Neuropsychology*, *1*(3), 183–211.

Shindell, M., McCaffrey, R. J., & Silk-Eglit, G. M. (2014, September). Three's a Crowd: The Impact of Third-Party Observers on Neuropsychological Exams. *Minority Trial Lawyer*.

Silver, C. H. (2000). Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *15*(4), 973–988.

Strauss, E., Sherman, E., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary* (Second). New York: Oxford University Press.

Stuss, D. T., Binns, M. a, Murphy, K. J., & Alexander, M. P. (2002). Dissociations within the anterior attentional system: effects of task complexity and irrelevant information on reaction time speed and accuracy. *Neuropsychology*, *16*(4), 500–513.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., … Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*(4), 388–402.

Swan, J., Robertson, M., & Evans, S. (2009). *Managing Clinical Research in the UK*. Coventry.

Towgood, K. J., Meuwese, J. D. I., Gilbert, S. J., Turner, M. S., & Burgess, P. W. (2009). Advantages of the multiple case series approach to the study of cognitive deficits in autism spectrum disorder. *Neuropsychologia*, *47*(13), 2981–8.

Volkmar, F. R., & Pauls, D. (2003). Autism. *The Lancet*, *362*, 1133–1141.

Wilson, B. a., Evans, J. J., Emslie, H., Alderman, N., & Burgess, P. (1998). The Development of an Ecologically Valid Test for Assessing Patients with a Dysexecutive Syndrome. *Neuropsychological Rehabilitation*, *8*(3), 213–228.

## Appendix I. Study Evaluation Criteria for Systematic Literature Review.

| Criteria | | Yes (2) | Partial (1) | No (0) | N/A |
|---|---|---|---|---|---|
| 1 [R] | Focus/question of study clearly identified | | | | |
| 2 [R] | Study design clearly described and appropriate | | | | |
| 3 [M] | Clinical and control groups selected in appropriate way | | | | |
| 4 [M] | Clinical group has a diagnosed autistic spectrum disorder meeting DSM criteria and at least one other screening measure to confirm diagnosis | | | | |
| 5 [M] | Autism symptomology screened in control group using appropriate method | | | | |
| 6 [M] | Were the clinical and control groups matched according to age, gender and FSIQ? (partial to include IQ matching based on only PIQ / VIQ) | | | | |
| 7 [M] | Outcome variable clearly defined and measured (at least one generativity task and it is clear what measure of dependent variable they have used) | | | | |
| 8 [M] | Study reports means along with standard deviations of the outcome variable | | | | |
| 9 [M] | Is the sample size appropriate / is there sufficient power (> …)? | | | | |
| 10 | Are analytic methods described and appropriate? | | | | |
| 11 [R] | Results reported in sufficient detail | | | | |
| 12 [R] | Conclusions supported by the results | | | | |

Total Score /24

Methodological Quality /16

Reporting Quality /8

# Appendix II. Eco-TED task script for Luria, Alternating Sequence, Pattern Drawing and Storytelling Tasks



**Luria hand movement task**
(Originally relating to item 8)

The child is asked to copy a sequence of hand movements initially demonstrated by the examiner. They should be positioned opposite you at the desk for this task. The hand movements should be demonstrated on the desk and the child asked to repeat them on the desk so that they can be easily observed by the examiner.

First ascertain the participant's dominant hand:

> ***Which hand do you throw a ball with? / Which hand do you write with?***

Introduce the subtest by telling the child:

> ***This time I am going to show you some movements that I would like you to copy. First I want you to watch what I do. Then when I say go I would like you to repeat what I did.***

The examiner should place both hands on the table in view of the child. The first trial begins with finger tapping.

NB for all items and trials, the demonstration can be repeated ONCE at the participants' request before they attempt to copy but this should be noted on the score sheet. Any observations (e.g. correcting a mistake) should also be noted on the score sheet.



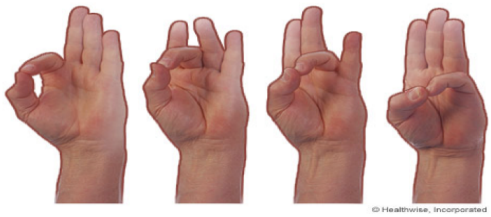
*Task Script*
*Version 8: February 2016*

**Trial 1**

> **Watch me.**

Start with the right hand, palm facing upwards. Begin by tapping each finger in order with your thumb. Start with the finger closest to the thumb. If the child looks away, give a prompt:

> **Remember I am going to ask you copy what I did so you need to watch carefully.**




When finished tell them:

> **Now it's your turn. Go.**

If the child demonstrates the sequence *correctly,* say the below and move on to **Item 1***:*

**That's it. Let's try another.**

If the child performs the trial *incorrectly:*

**That's not quite right. Watch me again. Remember I want you to copy exactly what I do once I say go.**

Repeat the sequence again on your right hand ensuring that the child is watching. If they manage to copy the sequence correctly, say the below and move on to **Item 1**:

**That's right, you've got it. Let's try another.**

If they perform the trial incorrectly, also continue to Item 1 and say:

**Let's try another.**





**Item 1**

**Now we are going to do something similar using our hands. I am going to show you a pattern of hand movements. First I want you to watch me. Then we will repeat the movements together.**

For this part of the test use the child's dominant hand.

With verbal prompts demonstrate the patterns leaving an equal gap between each movement (see Picture 1 below if unsure).

**Item 1(1)** *Palm – Palm – Palm (saying "**palm, palm, palm**" with every hand movement)*
**Item 1(2)** *Fist – Fist – Fist (saying "**fist, fist, fist**" with every hand movement)*
**Item 1(3)** *Edge – Edge – Edge (saying "**edge, edge, edge**" with every hand movement)*

If they do not repeat the item correctly the first time:

**That's not quite right - watch me again. It won't count unless your hands look just like mine.**

Repeat the pattern and prompts. Ensure that the participant makes the above hand movements with fingers straight and together (palm and edge) or in a clear fist. 2 points should be awarded if they repeat the pattern correctly the first time and 1 point if they repeat the pattern the second time.

Discontinue the task if participant scores 0 on any trial i.e. if they fail both trials. If they score 1 or 2 proceed to the next item. If item 1(3) is correct proceed to **item 2.** Say:

> ***Let's try some more. This time they are a little bit different. I will show you them more than once. When I have stopped I want you to copy what I did and repeat it the same number of times I do them.***

Repeat each pattern 3 times without verbal prompts. Movements should be demonstrated at a steady pace of one per second.

**Item 2:** Fist – Palm – Edge
**Item 3**: Fist – Edge – Palm
**Item 4:** Edge – Palm – Fist
**Item 5:** Edge – Fist – Palm
**Item 6:** Palm – Fist – Edge
**Item 7**: Palm – Edge - Fist

Score 1pt for the correct sequence and 1pt for the correct number of repetitions. Total = 2pt (see score sheet).

⇨**Discontinue rule:** Discontinue after 3 consecutive scores of 0.





Picture 1: Item 6 sequence of hand movements – palm-fist-edge

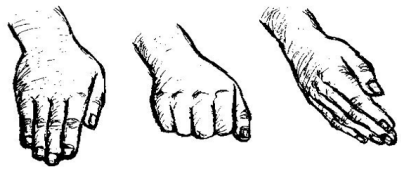


FIGURE 13.1. Luria's hand sequences (Luria, 1966).

**Alternating Sequence Task**
(Originally relating to item 8)

Items required:

- Alternating Sequence Stimulus Cards
- Alternating Sequence Response Sheet
- Black felt tip pen
- Stopwatch

For this task the participants will be asked to copy and complete a pattern of alternating shapes on a piece of paper. There will be three items and the patterns will be of increasing complexity. The child will have to draw the pattern from left to right of the response sheet without taking their pen off the paper.

**Item 1/Pattern 1:**
Place Pattern 1 stimulus card in front of the participant. Say the following:

> ***I am going to start by showing you a pattern on a piece of paper. You draw the pattern like this*** *(trace over the pattern so that the participant can see how the pattern is drawn without removing the pen from paper).* ***I want you to copy the pattern underneath exactly as you see it without taking your pen off of the paper. Start here*** *(point to left hand side of paper)* ***and finish here*** *(point to right side of paper).* ***Be as quick as you can. You can start when I say go.***





Ensure that the response page is positioned in front of the child, with the stimulus card above it, and place a black felt tip pen in front of the participant.

> ***You should draw the pattern here.*** *(*Point to the space for item 1/pattern 1 on the response sheet)
> ***Ready? Go.***

Start timing as soon as you have said "go". Stop timing once the participant reaches the end of the paper or once two minutes has elapsed. Record time on the score sheet and proceed to item 2.

A pattern correctly drawn from left to right of the page without removing pen from paper should be given a score of '2'. If the participant does not complete the pattern (score '0') or removes the pen from the paper (score '1'), give the following reminder:

> ***Remember you need to copy the pattern exactly as it is without taking your pen off of the paper. Let's try another one.***

If the participant is observed to perseverate (draws one of the shapes repeatedly) this should be recorded in the 'observations' section on the score sheet.

**Item 2/Pattern 2**
Place Pattern 2 stimulus card in front of the participant. Say the following:

***Here is the next pattern I would like you to draw. You should draw it in the space here*** *(point to correct place on response sheet)****. Remember draw it just like the picture without taking your pen off of the paper. You should do it as quickly as you can.***

***Go.***

Start timing as soon as you have said "go". Stop timing once the participant reaches the "finish" line or once two minutes as elapsed. Record time on the score sheet along with the awarded score for the pattern (0/1) and proceed to item 3.

**Item 3/Pattern 3**

Place Pattern 3 stimulus card in front of the participant. Say the following:

***Let's try another. Go.***

Start timing as soon as you have said "go". Stop timing once the participant reaches the "finish" line or once two minutes as elapsed. Record time on the score sheet along with the awarded score for the pattern (0/1). If the participant is unsure how the pattern is drawn, they should not be shown how to do it but should instead be encouraged to work it out themselves.

⇨**Discontinue Rule:** All three items should be attempted regardless of the participant's success on previous items.





**Pattern Drawing Task**

(Originally relating to item 8)

Items required:

- Pattern drawing stimulus pictures
- Response Booklet
- Black felt tip pen
- Stopwatch

The participant is shown a shape and told that they will copy the shape on a piece of paper. Initially they are given the chance to draw the shape by joining up numbered dots in order. They will then be asked to draw the shape a second time from memory. They will have to draw the shape without lifting their pen from the paper. There will be three different shapes to draw and the child will have three chances to draw each correctly (3 trials for each item). Begin with the following:

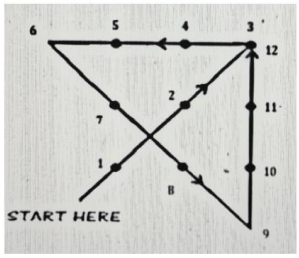
***I am going to start by showing you a picture of a shape. Then you are going to draw the shape by joining up some dots in the correct order without taking your pen off the paper. After this you are going to draw the shape again without the numbers to help you.***

**Item 1**

Show the picture of pattern 1 to the child. Place a black felt tip pen in front of the participant.

> ***This is the first shape you should draw. Now I want you to practice drawing the shape by joining up the dots in the correct order. Remember try to join them all without taking your pen off the paper.***

Examiner points to the 'demonstration 1' space in the response booklet and the participant is given the opportunity to join the dots (allow 120 seconds). Ensure that the child joins the dots in the correct order without removing their pen from the paper. The shape should be drawn as below:








If the child **correctly** draws the demonstration picture (joins all of the dots in the correct order without removing the pen from the paper) say:

> ***That's right. Now I want you to draw the shape again but this time there won't be any numbers. I want you to try and remember how you drew the shape last time*** *[remove the stimulus picture from view]****. You should draw it here*** *(turn to item 1, trial 1 in the response booklet)****. Remember your pen should stay on the paper. Be as quick as you can.***
>
> ***You can start when you are ready.***

Start timing as soon as the child places their pen to paper. Stop timing once the participant has drawn the shape or once one minute as elapsed. Record time on the response sheet and proceed to item 2.

If the child **incorrectly** draws the demonstration item turn to item 1, demonstration 2 in the response book where they will repeat the demonstration shape:

> ***That's not quite right. Let's try again. Draw the shape without lifting your pen from the paper. You can draw the shape by joining up the dots in the correct order for example 1,2,3…***
>
> ***Do you understand what you need to do?***

If the child draws the shape **correctly** proceed to the second part of the item (item 1, trial 2) where they are asked to draw the shape from memory:

> ***That's right. Now I want you to draw the shape again but this time there won't be any numbers. I want you to try and remember how you drew the shape last time . You should draw it here*** *(point to the space for item 1, trial 2 in the response booklet)****. Remember your pen should stay on the paper. Be as quick as you can. Start when you are ready.***

If **incorrect**, administer demonstration 3 and trial 3.

**Item 2**

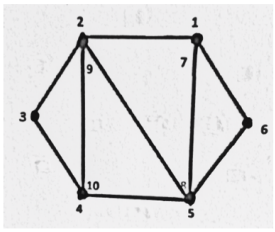
Show the picture of pattern 2 to the child. Say:

> ***This time we're going to try a different shape. This is the shape I'd like you to draw*** *(show participant Pattern 2)****. Let's try drawing it using the numbered dots*** *(turn to item 2 demonstration 1 in the response booklet).* ***Remember not to lift your pen from the paper.***
>
> ***You can start when you are ready.***

The pattern should be drawn as below:








If the child **correctly** draws the demonstration picture (joins all of the dots in the correct order without removing the pen from the paper) say:

> ***That's right. Now I want you to draw the shape again but this time there won't be any numbers. I want you to try and remember how you drew the shape last time*** *[remove the stimulus picture from view]****. You should draw it here*** *(turn to item 2, trial 1 in the response booklet)****. Remember your pen should stay on the paper. Be as quick as you can.***

Start timing as soon as you have said "go". Stop timing once the participant has drawn the shape or once one minute has elapsed. Record time on the response sheet and proceed to item 2.

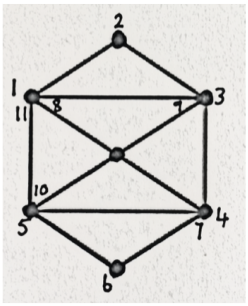
If **incorrect** administer subsequent demonstrations and trials.

**Item 3**

Turn to item 3, demonstration 1 in the response booklet and place pattern 3 in front of the child. Say:

> ***Here is a third shape. Practice drawing the shape by joining the dots.***

The pattern should be drawn as below:




> ***(NB the order of the dots doesn't match the stimulus booklet – EB amending in hard copy to go in testing pack)***





If the child **correctly** draws the demonstration picture (joins all of the dots in the correct order without removing the pen from the paper):

> ***That's right. Now I want you to draw the shape without the numbers*** *[remove the stimulus picture from view]****. You should draw it here*** *(turn to item 3, trial 1 in the response booklet)****. Remember keep your pen on the paper and be as quick as you can.***

Start timing as soon as the child puts pen to paper. Stop timing once the child completes the shape or once one minute as elapsed.

If **incorrect** administer subsequent demonstrations and trials.

**Scoring**

Points depend on the number of correctly drawn lines in the shape i.e. if the child was able to remember the first three lines of the shape but no more they would be awarded 3 points.

Item 1: max. 12 points.

Item 2: max. 10 points

Item 3: max. 11 points

For each item note the number of the trial where the participant draws the shape correctly (1-3). Mark 0 if the child was unable to draw the shape correctly in any of the three trials. Make a note if the shape is drawn correctly without removing pen from paper but in a different way to that shown in the demonstration.

**Story Telling Task**

(Relating to items 10 and 13: narrative coherence and storytelling)

Items required:

- Cartoon Strip (Story 1)
- Stimulus Cards (Story 2)
- Voice Recorder
- Scoring sheet

This task will test the participant's abilities to recall a story and to tell a story. Two short stories will be read to the child. Story 1/item 1 is accompanied by a comic strip. Story 2/item 2 is accompanied by four stimulus cards depicting important aspects of the story.

After the child has heard the first stories they will be asked to tell the story back to the researcher in as much detail as they can remember.

Item 3 involves the participant telling an account of their day so far, up to and including the testing session.





**Item 1/Story 1 "Lunchtime"**

Place the cartoon strip for story 1 in front of the participant.

> ***Now you're going to hear a story. This story is called "Lunchtime". I want you to listen carefully. When the story is finished, I am going to ask you to tell me what happened, trying not to leave anything out.***

Play the recording to the participant (text below for reference):

1. *It is lunchtime at school and James sits down to eat his lunch with the other children in his class. He opens his lunch box to find a sandwich, an apple and some raisins. James is disappointed. He really thought that his mum might have packed a chocolate bar for lunch today.*
2. *As he sits looking disapprovingly at the content of his lunchbox, Ben comes in to the class and puts his lunchbox on the table next to James. He asks James to look after his lunchbox whilst he goes to the toilet.*
3. *James sits watching over Ben's lunchbox but can't resist having a sneaky look to see what's inside. He checks that no one else is looking and lifts the lid a little. Straight away James spots a delicious chocolate bar. James mouth starts to water.*
4. *Sure that no one is looking, James quickly swaps the chocolate bar for his raisins and closes Ben's lunch box. He opens the chocolate bar and is feeling pleased with himself when Ben walks back in to the room.*
5. *Ben sits down and opens his lunch box. He lets out a big sigh which makes James jump. "What's wrong?" asks James. "I thought my mum might have given me a chocolate bar like you but she's given me raisins instead".*

6. *James turns red. "Why don't we share this one"? Ben smiles at James and takes a piece "Thanks James". James is left feeling less pleased with himself as the boys finish lunch and head into the playground.*

*Once the recording has finished, say:*

**Now tell me what happened in that story.**

***Start the voice recorder once instructions are given.***





**Item 2/Story 2 'Doris the Cat'**

Say:

***Now you're going to hear another story. This story is about 'Doris the Cat'. I want you to listen carefully. When the story is finished, I am going to ask you to tell me what happened in as much detail as you can.***

Play the recording to the participant (text below for reference). For each picture, place it in front of the participant at the correct time (indicated below) so that they can see it quite clearly, but ask them not to touch the pictures. ***Record*** any occasions when they break this rule [If they do pick up a picture, please ask them to put it back, and make sure each time that the picture is placed on top of the last one in a neat stack so that the previous picture is no longer visible].

***This is a story about Doris the cat*** *[show picture of Doris].*

***Doris is sitting at an open window, watching two birds in the garden outside. She is hiding behind a large plastic vase so that the birds can't see her. One bird, a robin, is trying to pull a worm out of the ground. Here is a picture of the robin*** *[show picture of robin].*

***While this robin is wrestling with the worm, another smaller bird is nearby, watching carefully for an opportunity to fly in and steal the worm if the Robin drops it.***

***As Doris watches the birds, she is gently swishing her tail about, left and right, left and right. Suddenly she feels a sharp pain in her tail that makes her jump up in surprise. She knocks over the plastic vase she was hiding behind and it falls to the ground with a terrible clatter. Luckily it does not break. Doris quickly looks round to see what had caused the pain and sees Elsie, her kitten, playfully nibbling on her tail*** *[show picture of kitten].*

***Doris looks back at the birds in the garden. When Doris knocked the vase over, the Robin was so surprised that it let go of the worm that it had finally managed to pull out of the ground. The Robin looks round to see Doris, and stands motionless watching the cat. As quick as a flash, the little grey bird that had been watching the robin swoops in and grabs the worm in its beak and flies off into a tree. Here is a picture of the tree*** *[show picture of tree]. **It has a long tall trunk so birds are safe from cats when they are high in the branches.***

***Back on the ground, the Robin looks back round to check on the worm it had dropped. But it is gone! The Robin looks all over the ground, but can't find it.***

***So the Robin flies onto an empty branch in the tall tree, so it could be safe from cats. As soon as it is comfortable on the branch, it sees the little bird sitting on a nearby branch, struggling with the worm. The worm is too big for the little grey bird to eat easily. While trying to swallow it, the little grey bird accidentally drops the worm and it lands near to the base of the tree. It lands on some leaves with a gentle rustling noise. The little bird quickly flies to the ground to try to pick it up.***

***But Doris has heard the worm fall onto the leaves, and now sees the little grey bird on the ground. She jumps out of the window and scampers playfully towards the little bird, as fast as she can. The little grey bird is frightened by the sight of Doris running over, and flies off into a nearby bush, leaving the worm where it had fallen, on the leaves.***





***Doris walks over to the bush to look where the little grey bird has gone, but she cannot see it in the thick bush.***

***The Robin has been watching this commotion from the tall tree. It sees the big worm left on the leaves, swoops down from the tree, picks it up, and returns to the branch, triumphant. Doris feels another nip on her tail, and looks round to see that Elsie the kitten has joined her to play.***

If the participants ask for pictures of things or protagonists in the story that are not available say:

***I'm sorry, I haven't got a picture of that.***

When you have read the story to the participant, turn the picture cards over so that the participant cannot see what is on them, and say:

**Now I would like you to tell me what happened in that story, in as much detail as you can. You can take as long as you like, but please tell me when you have told me as much as you can.**

***Start the voice recorder once instructions are given.***

**Item 3**

This task is **to be administered at the end of the session** (after BADS-C and WASI). The participant will be asked to give an account of the session. Say:

> ***Now I want you to tell me everything that has happened since I/you got here** [as appropriate].*

***Start the voice recorder once instructions are given.*** The instructions may be repeated as requested. Nodding and general statements of interest should be given by the examiner to encourage the participant.

## Appendix III. Trainee's individual contributions to the research project

This project was undertaken alongside another Trainee studying for the Clinical Psychology Doctorate (UCL). During the initial stages of task design, we worked alongside our supervisor's Dr William Mandy and Professor Paul Burgess to identify those items of the 3Di for which we would design tasks. Once we had a joint consensus on which ideas to pursue I took sole responsibility for developing the four tasks outlined in this thesis paper. Throughout the project, we had research team meetings to discuss feasibility of the project and how our tasks might combine to create the assessment battery (the Eco-TED). Joint research ethics was sought for the project and data collection was carried out by one of the two researchers or one of three research assistants. The write-up of this paper was conducted independently by Jodie Pullinger.

# Appendix IV. Scoring for storytelling task

**Item / Story 1 – 'Lunchtime' Scoring Sheet**

| ***Event*** | | *Tick if included in the account* |
|---|---|---|
| 1 | James sits down to eat lunch with other children in his class | |
| 2 | James opens lunch box to find a sandwich, an apple and some raisins | |
| 3 | James is disappointed because his mother has not packed a chocolate bar / looks disapprovingly at the content of his lunchbox | |
| 4 | Ben comes in to the class and puts his lunchbox on the table next to James | |
| 5 | Ben asks James to look after his lunchbox whilst he goes to the toilet. | |
| 6 | James watches over Ben's lunchbox and sneaks a look inside | |
| 7 | After checking that no one is looking he lifts the lid and spots the chocolate bar / his mouth starts to water | |
| 8 | Swaps the chocolate bar for the raisins and closes Ben's lunch box | |
| 9 | Ben walks back into room / sits down and opens lunch box | |
| 10 | Ben lets out a big sigh which causes James to ask "what's wrong?" | |
| 11 | Ben replies "I thought my mum might have given me a chocolate bar like you but she's given me raisins instead". | |
| 12 | James turns red | |
| 13 | James offers to share the chocolate bar | |
| 14 | James feels bad | |
| 15 | The boys finish lunch and head into the playground. | |
| | Total | /15 |

| **Order of Events** | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | Total | /15 | |

| | Total | Notes |
|---|---|---|
| Time taken to recount story | | |
| Number prompts required to continue story | | |
| Number of pauses over 2 seconds long | | |
| Number of ToM / mentalising words in account | | |
| Number new elements (confabulation) | | |

| | | |
|---|---|---|
| Does the child indicate when they have finished retelling the story? | 0<br>No | 1<br>Yes |
| Does the child use inappropriate speech? | 0<br>No | 1<br>Yes |

| | Total | Notes |
|---|---|---|
| How many times does the child repeat any elements of the story? | | |
| Number of incomplete sentences | | |
| Number of times child refers to a person / character without it being clear which they are referring to (or confusing names) | | |
| Number of times the child confuses gender | | |
| Eye contact | | 0<br>No<br>1<br>Yes |

**Item / Story 2 – 'Doris the Cat' Scoring Sheet**

| ***Event*** | | *Tick if included in the account* |
|---|---|---|
| 1 | Doris is sitting at an open window watching two birds in the garden outside/. | |
| 2 | One of the birds, a robin, is trying to pull a worm out of the ground/. | |
| 3 | Meanwhile a small grey bird sits watching nearby ready to steal the worm if the Robin drops it/. | |
| 4 | As Doris watches the birds, she is gently swishing her tail about, left and right, left and right/. | |
| 5 | Suddenly she feels a sharp pain in her tail that makes her jump up in surprise/. | |
| 6 | She knocks over the plastic vase she was hiding behind which falls to the ground with a terrible clatter but does not break. | |
| 7 | Doris quickly looks round to find Elsie, her kitten, playfully nibbling on her tail | |
| 8 | The Robin was so surprised by the sound of the vase that it let go of the worm. | |
| 9 | As quick as a flash, the little grey bird swoops in and grabs the worm in its beak and flies off into a tree. | |
| 10 | Back on the ground, the Robin looks back round to check on the worm it had dropped. But it is gone! The Robin looks all over the ground, but can't find it. | |
| 11 | He looks up to see the little bird sitting on a nearby branch, struggling with the worm. The worm is too big for the little grey bird to eat easily. | |
| 12 | While trying to swallow it, the little grey bird accidentally drops the worm and it lands back on the ground. | |
| 13 | As the little grey bird swoops back down to collect the worm Doris jumps out of the window and scampers playfully towards the little bird, as fast as she can. | |
| 14 | The little grey bird is frightened by the sight of Doris running over, and flies off into a nearby bush. | |
| 15 | The Robin has watched the commotion from the tall tree and sees the worm on the leaves. It swoops down from the tree and picks it up, returning to the branch, triumphant. | |
| | Total | /15 |

| **Order of Events** | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | Total | /15 | |

| | Total | Notes |
|---|---|---|
| Time taken to recount story | | |
| Number prompts required to continue story | | |
| Number of pauses over 2 seconds long | | |
| Number of ToM / mentalising words in account | | |
| Number new elements (confabulation) | | |

| | | |
|---|---|---|
| Does the child indicate when they have finished retelling the story? | 0<br>No | 1<br>Yes |
| Does the child use inappropriate speech? | 0<br>No | 1<br>Yes |

| | Total | Notes | | |
|---|---|---|---|---|
| How many times does the child repeat any elements of the story? | | | | |
| Number of incomplete sentences | | | | |
| Number of times child refers to a person / character without it being clear which they are referring to (or confusing names) | | | | |
| Number of times the child confuses gender | | | | |
| Number of seen items | | | | |
| Number of unseen items | | | | |
| Eye contact | | | 0<br>No | 1<br>Yes |

**Item 3 – Retelling of the session**

| *Event* | | *Tick if included in the account* |
|---|---|---|
| 1 | Introduction / coming into session or discussion of what they have done earlier in day | |
| 2 | Information about study / being told what the session will involve | |
| 3 | Filling in consent form | |
| 4 | School Bag Task – recall should include being asked to pack imaginary school bag **or** reference to at least one specific lesson | |
| 5 | Lego Task - should include reference to having to search for Lego pieces | |
| 6 | Lego Task – searching for favourite animal | |
| 7 | Luria Hand Task (e.g. having to copy hand movements) | |
| 8 | Alternating Sequence Task (e.g. having to drawn patterns on paper / copying patterns) | |
| 9 | Pattern Drawing Task (e.g. having to join dots / remember a pattern / not remove hand from paper) | |
| 10 | Story Telling Task 1 Lunchtime - does not need to recall story again | |
| 11 | Story Telling Task 2 Doris the Cat – does not need to recall story again | |
| 12 | BADS 6 Part Test (e.g. sorting through things, having to do sums, write names of pictures etc) | |
| 13 | BADS Zoo Map Test (e.g. plot the way around a zoo) | |
| 14 | WASI Vocab (e.g. describe the meaning of the word) FOR THOSE COMPLETING WASI ONLY | |
| 15 | WASI Matrix Reasoning (e.g. pick the correct pattern) FOR THOSE COMPLETING WASI ONLY | |
| 16 | Ending – makes some reference to ending the session / finishing all the tasks | |
| | Total | WASI /16<br>No WASI /14 |

| **Order of Events** | | |
|---|---|---|
| | | |
| Total | WASI /16<br>No WASI /14 | |

| | Total | Notes |
|---|---|---|
| Time taken to recount story | | |
| Number prompts required to continue story | | |
| Number of pauses over 2 seconds long | | |
| Number of ToM / mentalising words in account | | |
| Number new elements (confabulation) | | |

| | |
|---|---|
| Does the child indicate when they have finished retelling the story? | 0 No 1 Yes |
| Does the child use inappropriate speech? | 0 No 1 Yes |

| | Total | Notes |
|---|---|---|
| How many times does the child repeat any elements of the story? | | |
| Number of incomplete sentences | | |
| Number of times child refers to a person / character without it being clear which they are referring to (or confusing names) | | |
| Number of times the child confuses gender | | |
| Eye contact | | 0 No 1 Yes |

# Appendix V. Strengths and Difficulties Questionnaire

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

Do you have any other comments or concerns?

# Appendix VI. Research Ethics – Letter confirming favourable ethical opinion

NHS
Health Research Authority
Research Ethics Service

**London - Westminster Research Ethics Committee**
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0207 104 8012

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, nrescommittee.london-westminster@nhs.net.

**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 (catherineblewett@nhs.net), 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:

| *Document* | *Version* | *Date* |
|---|---|---|
| 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

[redacted]

**Dr Alan Ruben**
**Chair**

Email:nrescommittee.london-westminster@nhs.net

Enclosures: "After ethical review – guidance for researchers"

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

# Appendix VII. Initial Contact Letter: Parents/caregivers of children with ASD

Version 2.0 11/03/2015

Great Ormond Street Hospital for Children NHS Foundation Trust

[redacted]
Department of Clinical, Health and Educational Psychology
University College London
**1-19 TORRINGTON PLACE**, LONDON, WC1E 7HB
**Tel:** [redacted]

**Email:** [redacted]

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 can 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 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 also 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.

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] (the researchers) or return the enclosed consent form so that we can contact you. Our details can be found at the top of this page.

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]

*(Researcher)* *(Researcher)* *(Chief Investigator)*

# Appendix VIII. Information & Consent Form: Parents/Caregivers of children with ASD

Version 7.0 08/09/2015

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL

**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 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 077xxx or email [redacted]. In the case of a complaint, please contact [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.*

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL
Department of Clinical, Health and Educational Psychology
University College London
**1-19 TORRINGTON PLACE**, LONDON, WC1E 7HB
**Tel:** [redacted]
**Email:** [redacted]

**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 a reason.

☐ 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: _____

# Appendix IX. Initial Contact Letter: Parents/caregivers of typically developing children

Great Ormond Street Hospital for Children NHS Foundation Trust

[redacted]
Department of Clinical, Health and Educational Psychology
University College London
1-19 TORRINGTON PLACE, LONDON, WC1E 7HB
Tel: [redacted]

Email [redacted]

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,

[redacted] *(Researcher)*

[redacted] *(Researcher)*

[redacted] *(Chief Investigator)*

# Appendix X. Information & Consent Form: Parents/Caregivers of typically developing children

Version 6.0 08/09/2015

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL

**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 Eleonore or Jodie (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 077xxx 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.*

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL
[redacted]
Department of Clinical, Health and Educational Psychology
University College London
**1-19 TORRINGTON PLACE**, LONDON, WC1E 7HB
**Tel:** 07794xxxxxx
Email [redacted]

**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 a reason.

☐ 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: _____

# Appendix XI. Information & Assent Form: Child participants

Version 2.0 10/04/15

## INFORMATION SHEET FOR CHILDREN & YOUNG PEOPLE

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL

### 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 Eleonore or Jodie 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 Eleonore or Jodie.

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL

☐ 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, Eleonore or Jodie

☐ 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 XII. Letter to Head Teacher

Great Ormond Street Hospital for Children NHS Foundation Trust

UCL [redacted]
Department of Clinical, Health and Educational Psychology
University College London
**1-19 TORRINGTON PLACE**, LONDON, WC1E 7HB
**Tel:** [redacted]

**Email:** [redacted]

Invitation for your school to take part in research

Dear *(Insert Name Head teacher)*

Your details have been passed on to us as you have kindly said that *(insert name of school)* would be open to and interested in taking part in the doctoral research project that we are conducting.

The project's aim is to validate a newly developed measure of executive functioning for children with Autistic Spectrum Disorder (ASD). We are looking to recruit a control group of around 40 children from London schools who do not have a diagnosis of ASD. We are therefore very grateful to you for considering to support this research.

The study will involve no more than two hours of the children's time and can be conducted at their home, school or at Great Ormond Street Hospital. 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. Parents will also be asked to complete some questionnaires about their 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 the child will take part in this second visit which should take no longer than one hour. As a thank you the child will receive a £5 voucher and will also be entered into a draw to win one of two £50 vouchers.

We have enclosed our participant information sheets so that you have a better idea of what we will be asking of parents and children. We would be happy to come and meet with you to discuss in more detail or to answer any questions that you might have.

Once again, thank you for your time and for considering taking part in our research.

Yours Sincerely,

[redacted] *(Researcher)*

[redacted] *(Researcher)*

[redacted] *(Chief Investigator)*