UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL AND HUMAN SCIENCES

Psychology

Learning to wait: the development and initial evaluation of a training intervention designed to help impulsive preschool children at risk for ADHD learn to wait for rewards

by

Pavlina Markomichali

Thesis for the degree of Doctor of Philosophy

February 2015

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL AND HUMAN SCIENCES

Psychology

Thesis for the degree of Doctor of Philosophy

LEARNING TO WAIT: THE DEVELOPMENT AND INITIAL EVALUATION OF A TRAINING INTERVENTION DESIGNED TO HELP IMPULSIVE PRESCHOOL CHILDREN AT RISK FOR ADHD LEARN TO WAIT FOR REWARDS

Pavlina Markomichali

Recent evidence suggests that early intervention can produce considerable improvements in the neurocognitive performance of young children at risk for ADHD. Acknowledging the role of motivational dysfunction in ADHD, the current thesis seeks to extend the scope of this approach by developing a training program that targets sub-optimal motivational processes in very young children. Specifically, it developed and evaluated a delay training paradigm aimed at improving young children's ability to wait for delayed rewards. The results of a small scale randomized controlled trial indicated that impulsive children's ability to wait for delayed rewards significantly improved during the training sessions, but that these gains did not generalize to improvements in untrained measures of delay tolerance. These findings help address the question of whether this ability is amenable to practice and have potential implications for the management of delay-related difficulties in clinical and educational contexts.

Thesis Outline

Chapter One provides the background to the clinical characteristics and neuropsychological alterations associated with attention deficit hyperactivity disorder (ADHD).

Chapter Two examines the dominant treatment model for ADHD and lays out the theoretical rationale for early and preventive intervention as a new direction to ADHD treatment. It also reviews the emerging evidence base for early intervention in ADHD.

Chapter Three makes the case for impulsive choice as a target for ADHD training interventions by providing an overview of findings drawn from different fields of psychological research showing that young children's ability to wait for delayed rewards is amenable to practice.

Chapter Four describes the development and the validation of two measures of delay tolerance which were developed or adapted for a preschool aged population and specifically for this thesis: a developmentally sensitive computerized choice delay task and the parent/teacher version of the Quick Delay Questionnaire, a brief rating scale of children's delay related difficulties.

Chapter Five describes the development of a novel delay training paradigm, the Waiting Game task, and presents findings from two feasibility studies demonstrating its training efficiency.

Chapter Six presents the focal study of the thesis: a small scale randomised controlled trial designed to evaluate the training efficiency and the training effects associated with the adapted Waiting Game training task in a sample of preschool aged children presenting with high levels of delay related difficulties.

Chapter Seven brings together the findings from the above studies in order to evaluate the potential theoretical and practical implications of the Waiting Game trials for the management of delay related difficulties in clinical and educational contexts.

Table of Contents

List of Tables

List of Figures

DECLARATION OF AUTHORSHIP

I, PAVLINA MARKOMICHALI, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Title of Thesis: "Learning to wait: the development and initial evaluation of a training intervention designed to help impulsive preschool children at risk for ADHD learn to wait for rewards"

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Markomichali, P., Donnelly, N. & Sonuga Barke, E.J.S. (2009). Cognitive training for attention, inhibition and working memory deficits: a potential treatment for ADHD? *Advances in ADHD*, 3(3), 89-98.

Signed:...

Date:..

Acknowledgements

With many thanks to

Prof. Edmund Sonuga-Barke for his continuous, tireless and inspiring supervision over all these years. A special mention should be made of our recurrent discussions on the nature of delay inescapability which over time helped shape the theoretical backbone of this thesis.

Dr. Jana Kreppner and Prof. Nick Donnelly for their unique contributions to various stages of this work.

Dr. Matt Jones for his invaluable technical support. His software development skills are out of this world.

My friends and colleagues for being so encouraging and never refusing to answer any of my countless questions. I would especially like to thank Dr. Evi Bitsakou, Dr. Suzannah Helps, Dr. Lamprini Psychogiou, Dr. Hanna Kovshoff, Dr. Helen Richards, Dr. Flis Bishop, Dr. Panagiota Andreou, Johanna Koerting, Chia Fen Hsu, Nick Benikos and Nadia Peppa.

My great friend and amazing young researcher Dr. Angeliki Bogosian not only for her emotional and practical support but also for being such a great role model.

My family and support network. I would particularly like to thank my daughter for all her patience. The amount of time required to complete this work almost precisely corresponded to the amount of time I have had to spend away from her in her four years of life. I very much hope there will come a point in the future when she finds all the anguish (on both our parts) was worthwhile.

All the children and the teachers who took part in these studies. Without their contribution none of this would have been possible.

This research was funded by the Institute of Disorders of Impulse and Attention, Developmental Brain-Behaviour Lab, Academic Unit of Psychology, University of Southampton.

Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AMA	Aimless motor activity
ANOVA	Analysis of Variance
BRIEF-P	Behavior Rating Inventory for Executive Function-Preschool Version
CBCL	Child Behavior Checklist
CD	Conduct Disorder
CDT	Cookie Delay Task
CSRP	Chicago School Readiness Project
D	Distracted Attention
DA	Delay aversion
DD	Delay discounting
DeFT	Delay Frustration Task
DG	Delay of gratification
DRT	Delay Reaction Time task
DSM	Diagnostic and Statistical Manual of mental disorders
DT	Delay Time
EF	Executive function
ENGAGE	Enhancing Neurocognitive Growth with the Aid of Games and Exercise
ETAM	Executive Training of Attention and Metacognition
IDIR	impulsive drive for immediate reward
IY	Incredible Years
LA	Looks up or away
LL rewards	Larger, later rewards
M	Mean
MANOVA	Multivariate Analysis of Variance
mDT	mean Delay Time
MIDA	Maudsley Index of Delay Aversion task
MTA study	Multimodal Treatment of ADHD study
NFPP	New Forrest Parenting Package
ODD	Oppositional Defiant Disorder
OS	Out of seat
PATHS	Promoting Alternative Thinking Strategies
PATS study	the Preschool ADHD Treatment Study
PS	Play with stickers
PT	parent training
QDQ	Quick Delay Questionnaire
RCT	randomised controlled trial
RMA	Restless motor activity
ROMA	Reward-oriented motor activity
SAT	Scholastic Aptitude Test

SD	Standard Deviation
SS rewards	Smaller, sooner rewards
T1	Time 1; first measurement point
T2	Time 2; second measurement point
TEAMS	Training Executive, Attention and Motor Skills
V	Verbalization
WG	Waiting Game
WWT	Watch-and-Wait Task
Year R	Reception Year

Chapter 1: Clinical characteristics and the neuropsychology of ADHD

1.1. Aim of the chapter

The aim of this chapter is to provide the background to the clinical characteristics and neuropsychological alterations associated with ADHD.

1.1. Clinical characteristics of ADHD

1.1.1. Epidemiology

Attention deficit hyperactivity disorder is an extreme and age inappropriate expression of childhood hyperactivity. It is the most prevalent disorder of childhood, affecting 5-8% of the general childhood population (American Academy of Pediatrics, 2011; American Psychiatric Association, 2013). Different prevalence rates have often been reported for different countries, but these are largely thought to be due to the use of narrower diagnostic definitions of the disorder (Swanson et al., 1998). A world prevalence study has estimated the worldwide-pooled prevalence of ADHD at 5.29%, but has also noted the significant variability associated with this estimate (Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007). Boys with ADHD outnumber girls by a ratio ranging from 2:1 to 9:1 but the gender difference decreases with age and is less obvious in community based, as opposed to referred, samples possibly indicating that ADHD is less disruptive in girls than in boys (Root & Resnick, 2003).

1.1.2. Core symptoms

Clinically, ADHD is characterized by pervasive and developmentally inappropriate symptoms of inattention, hyperactivity and impulsivity. To receive a diagnosis according to the fifth edition of the diagnostic and statistical manual of mental disorders (DSM 5), children must have six of nine symptoms of inattention (e.g. distractibility, poor concentration) or six of nine symptoms of hyperactivity (e.g. fidgeting, restlessness) and/or impulsivity (e.g. acting out without thinking, frequently interrupting others). If at least six symptoms of either category (inattention or hyperactivity/impulsivity) are present, then a diagnosis of predominantly inattentive or predominantly hyperactive/impulsive presentation can be reached. If at least 6 symptoms of both categories are present, then a diagnosis of combined presentation can be reached. But for a diagnosis to be applicable, these symptoms must have been present before the age of 12 and they must interfere with or reduce the quality of functioning in important life domains (e.g. social, academic or occupational functioning)(American Psychiatric Association, 2013)

1.1.3. Impact and other areas of impairment

The impact of this type of symptomatology on a child's life is often persistent and pervasive. For example, children diagnosed with ADHD often under perform at school: they achieve lower scores at standardised tests, make more use of educational support services and leave school with fewer qualifications compared to their non-ADHD counterparts (Loe & Feldman, 2007). They are more likely to have fewer friendships and to have difficult relationships with family members and/or their peers (Nijmeijer et al., 2008). In preschool years, they experience more motor coordination problems and tend to have more accidents (Lahey et al., 1998). As adolescents, they are at greater risk of engaging in antisocial behaviours or developing substance abuse problems, presumably as a result of their hyperactive and impulsive behaviours (Molina et al., 2013). Overall, the pattern of impairment changes over the years with different aspects of the disorder being more prominent at different life stages, but for some individuals the impact of ADHD can persist into adulthood and lead to life-long impairment (Harpin, 2005).

1.1.4. Comorbidity

It is widely accepted that the ADHD cluster of symptoms, albeit dissociable from other common childhood disorders, still tends to be highly comorbid (i.e. to frequently co-exist) with several neuropsychiatric conditions (Taylor & Sonuga-Barke, 2008). Perhaps the most notable association is with Oppositional Defiant Disorder (ODD), with as many as 50-60% of ADHD children also meeting diagnostic criteria for ODD (Gillberg et al., 2004). Other frequently co-existing disorders include Conduct Disorder (CD), depression, anxiety and mood disorders, bipolar disorder and substance abuse disorders, among others. It thus seems that, in clinical terms at least, it is very rarely the case that a diagnosis of ADHD is not accompanied by one or more additional DSM diagnoses. This degree of overlap is one of the most important characteristics of ADHD as it affects long term outcome and complicates decisions about intervention options (Gillberg, et al., 2004).

1.1.5. Developmental trajectory

Most commonly children are diagnosed with ADHD in middle childhood but onset is typically in the preschool years (three to six years of age) with the number of children receiving diagnoses while still of a preschool age rapidly increasing (Sonuga-Barke, Thompson, Abikoff, Klein, & Brotman, 2006). By the time children turn four, a diagnosis of ADHD is likely to persist into the school years (Lahey et al., 2004) and into adolescence (Lee et al., 2008). However, longitudinal studies exploring the continuity of symptoms and developmental outcomes across the life span have unveiled a more complex picture. On the one hand, symptoms of inattention, hyperactivity and impulsivity have been found to cluster together in a way that differentiates them from other types of oppositional or emotional problems, both in childhood (Rohde et al., 2001) and in earlier years (Sonuga-Barke, Thompson, Stevenson, & Viney, 1997). This confirms the validity of the concept of ADHD. On the other hand, it is also evident that some of the core ADHD behaviours manifest themselves and affect individuals in different ways at different ages (for example, as children get older they may become less hyperactive and more inattentive or risk prone, Klein & Mannuzza, 1991). In addition, there is a considerable degree of developmental discontinuity and remission. For instance, the proportion of children who do not meet diagnostic criteria after moving from preschool to school is around 20% (Lee, et al., 2008) and this figure is even larger in the case of adolescents entering adulthood (Mannuzza, Klein, Bessler, Malloy, & LaPadula, 1998).

1.1.6. Preschool ADHD

These challenges are particularly relevant in the case of preschool ADHD, often questioning the validity of the very construct of preschool ADHD. These concerns are further compounded by the difficulty of reliably distinguishing age related levels of high energy and activity, which some adults may find difficult to manage, from clinically relevant symptoms. In response to these concerns, Sonuga-Barke and colleagues have convincingly argued that preschool and school-aged ADHD should be seen as equivalent, particularly if ADHD symptoms are described as an extreme expression along a continuous dimension of behaviour as opposed to a distinct disorder-category (Coghill & Sonuga-Barke, 2012; Sonuga-Barke, Auerbach, Campbell, Daley, & Thompson, 2005; Sonuga-Barke, Daley, Thomson, & Swanson, 2003). According to this conceptualisation, despite the heterogeneity within the ADHD cluster and the overlap with other related childhood disorders, it is still of considerable clinical utility to view ADHD as an early-onset dimensional disorder. As already mentioned above, there is evidence supporting the clustering of ADHD symptoms (a six year follow-up of the preschool ADHD treatment study (PATS) recently confirmed that children with moderate to severe ADHD continued to meet criteria for ADHD at the 3-, 4-, and 6- year follow up, Riddle et al., 2013), and their distinctiveness from other behaviour problems even at preschool years (Sonuga-Barke, et al., 1997). Moreover, this early clustering of symptoms is associated with the same pattern of impairment, developmental risk and underlying neuropsychological functioning as school-aged ADHD, further supporting the utility of the construct (Dalen, Sonuga-Barke, Hall, & Remington, 2004; Daley, Jones, Hutchings, & Thompson, 2009).

But there is one caveat: the extension of the diagnostic category of ADHD into the preschool years can only be of value if age appropriate diagnostic items, definitions and thresholds are established. Relevant guidelines have been proposed. For example, Campbell (2002) proposed a differentiation of behaviours that occur with higher frequency, severity and duration, and that are present across settings (e.g. home, childcare) and relationships (e.g. parents, childcare providers), compared to behaviours that arise as a result of difficult developmental transitions (e.g. birth of a sibling or entry into childcare). The recently updated fifth edition of the DSM has acknowledged the need for developmentally appropriate diagnostic definitions of ADHD by including multiple examples of how the core symptoms present across different ages (American Psychiatric Association, 2013). But undoubtedly there is more work to be done before age-appropriate diagnostic criteria for preschool ADHD are in place.

1.2. The neuropsychology of ADHD

A clinical diagnosis is necessarily based on the assessment of observable behaviours and of the impact these behaviours have on an individual's functioning in daily life. But research in the field of child (and adult) psychopathology, has long sought to move beyond the behavioural level and to describe mental disorders in terms of disruption of function at a neuropsychological and/or neurobiological level of analysis. Particularly in the field of ADHD, research into neuropsychological alterations associated with the disorder has highlighted a marked heterogeneity characteristic of the disorder at many levels of analysis, e.g. developmental, psychopathophysiological and clinical (Fair, Bathula, Nikolas, & Nigg, 2012; Sjowall, Roth, Lindqvist, & Thorell, 2013; Sjowall & Thorell, 2014). The heterogeneous character of ADHD has in turn critical implications for the development of intervention strategies.

1.2.1. Executive dysfunction

In line with the dominant "bio-medical" model of mental disorders, early models of ADHD sought to identify the site of a single, core underlying dysfunction that "causes" the clinical phenotype of the disorder (Sonuga-Barke, 2005). For example, according to what has been the dominant model of ADHD until recent years, executive dysfunction was seen as the primary neuro-cognitive deficit responsible for ADHD. Although many variations of this model exist, in its boldest form it regards deficient inhibitory control as the core deficit which secondarily disrupts other executive function (EF) processes and gives rise to the broader pattern of executive impairment which is often associated with ADHD (Barkley, 1997b). These deficits are postulated to be grounded in dysfunction in dorsal frontostriatal neural circuits which are modulated by catecholamines, such as dopamine and norepinephrine (see Durston, van Belle, & de Zeeuw, 2011 for a recent explication of the neurobiological basis of ADHD). This conceptualisation of the disorder has been very influential but is only partially supported by empirical evidence.

Even though children with ADHD consistently exhibit suboptimal performance on EF measures at a group level when compared to non-ADHD controls (Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005), EF deficits are not identifiable in all individuals with ADHD. A key meta-analysis which has looked into EF deficits assessed by a multitude of tasks in a large ADHD sample (over 6700 children) illustrates this point (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). Even the most discriminating EF tasks, could only identify around 50% of children with ADHD as impaired, when the cut-off for impairment was set as the 90th percentile of controls' scores. And although around 80% of children with ADHD were found to be impaired in at least one EF task, the same was true for almost half of the control children as well. These findings suggest that executive dysfunction is neither specific to ADHD nor is it the sole etiologic pathway to the emergence of the disorder.

1.2.2. Context-dependent deficits

It is an interesting fact that cognition in ADHD is inconsistent and highly variable from context to context (Castellanos, Sonuga-Barke, Milham, & Tannock, 2006). Models focusing on the dynamic and context-dependent nature of the disorder have implicated more deep-seated suboptimal motivational processes and state regulation deficits as alternative pathophysiological bases for the disorder.

1.2.2.1. Motivational dysfunction. Children with ADHD have been consistently found to favour small immediate to large delayed rewards in simple choice tasks when compared to controls (Bitsakou, Psychogiou, Thompson, & Sonuga-Barke, 2009; Sagvolden, Johansen, Aase, & Russell, 2005; Sonuga-Barke, Taylor, Sembi, & Smith, 1992; Tripp & Alsop, 2001). In other words, ADHD children prefer rewards that help them minimise time on task, while control children typically opt for rewards that help them maximise their total gains. This behaviour has been explained as resulting from failures in the signalling of delayed rewards (Sagvolden, et al., 2005) or from aversion to delay (Sonuga-Barke, 2002, 2003; Sonuga-Barke, Taylor, Sembi, et al., 1992). It is also associated with altered patterns of activation in brain reward systems, and particularly the ventral striatum, during reward anticipation (Durston, et al., 2011; Edel et al., 2013). Interestingly, this preference for immediacy over delay has proven to be an effective motivational marker for ADHD as it can differentiate children with ADHD from controls to a satisfactory degree, with pooled effect sizes ranging from 0.57 to 0.71, depending on the type of task used (Willcutt, Sonuga-Barke, Nigg, & Sergeant, 2008). These effect sizes are comparable to those seen for executive function measures.

1.2.2.2. State-regulation models. According to the cognitive-energetic model (Sergeant, 2000), cognitive functioning is mediated by an individual's state of physiological arousal, activation level (i.e. degree of readiness for action) and effort allocation. When applied to the case of ADHD cognition, this model can explain the impact of event rate (presentation rate of stimuli) on cognitive performance and successfully predict the deterioration of performance during periods of under-activation -typically operationalized as low event rate conditions- which is a consistent finding in the ADHD literature (van der Meere, 2005). The exact brain circuitry underpinning state regulation deficits in ADHD is not easy to define, but there is fast accumulating evidence supporting the disruption of low-frequency brain connectivity in ADHD. This type of disruption may explain the existence of periodic attention lapses interfering with task-specific and goal-directed activity and may be related to the dynamic and variable nature of cognition in ADHD (Sonuga-Barke, 2013).

1.2.2.3. Timing deficits. Children with ADHD have marked difficulties when asked to process time related information (Toplak, Dockstader, & Tannock, 2006). For example, researchers have demonstrated that children with ADHD display poorer performance on tasks used to measure time perception. Examples include tasks involving the comparison of two brief intervals similar in duration to determine which is longer or shorter (duration discrimination tasks), tasks requiring the participant to reproduce the duration of a specified interval (duration production tasks), tasks requiring the tapping of fingers at a specified pace, and others (Huang et al., 2012; Rubia, Noorloos, Smith, Gunning, & Sergeant, 2003; Sonuga-Barke, Bitsakou, & Thompson, 2010). Clearly the processing of temporal information is a complex and multi-stage cognitive skill involving various component processes and engaging different brain regions. Research is only now starting to disambiguate some of these constructs but already there is evidence indicating that temporal information deficits may represent a distinct type of neuropsychological deficit in ADHD, dissociable from both executive and motivational dysfunction (Coghill, Seth, & Matthews, 2014; Sonuga-Barke, Bitsakou, et al., 2010).

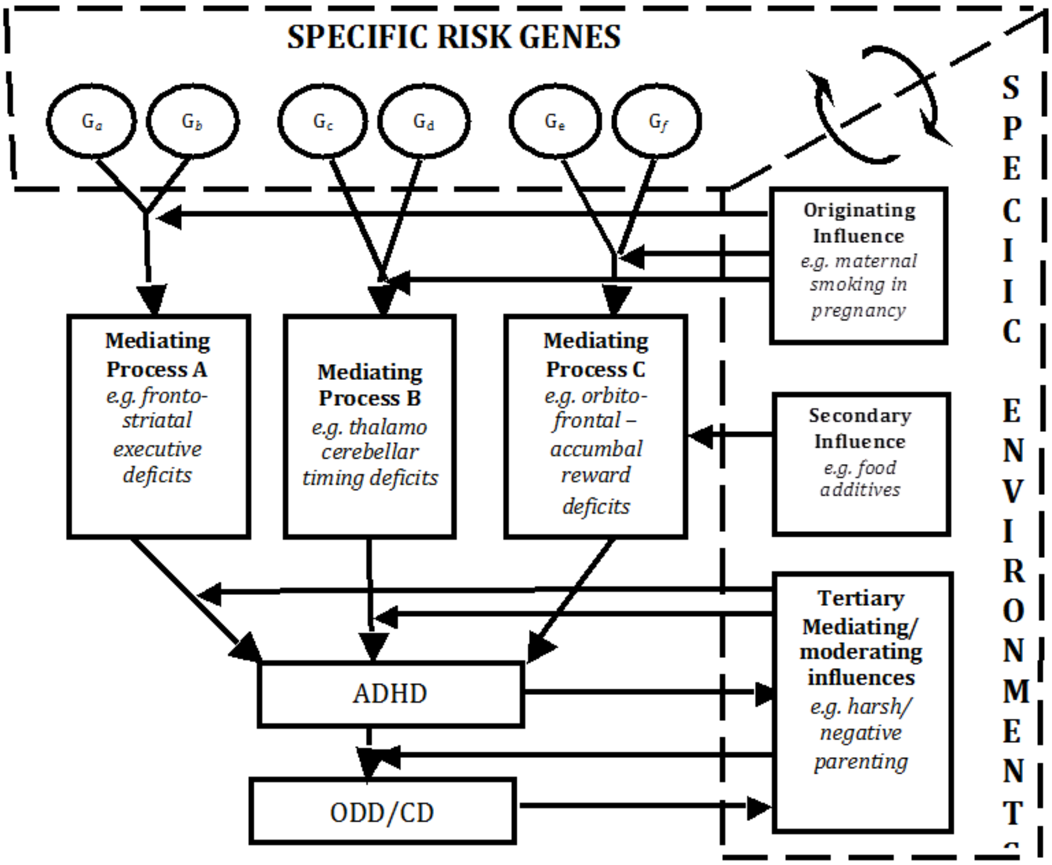
1.3. The challenge of heterogeneity and the multiple-pathway model

To summarise so far, the study of the neuropsychology of ADHD has pinpointed several important challenges to the older, medical conceptualisation of the disorder: a. the existence of a dimensional aspect to ADHD exemplified by its changing nature across the life span; b. the degree of overlap between ADHD and other disorders, as evidenced by the unusually high rates of comorbidity and by the lack of specificity to ADHD of even the most characteristic neuropsychological weaknesses associated with the disorder and c. the marked heterogeneity of children diagnosed with ADHD at a clinical, etiological and neuropsychological level.

The third challenge in particular, i.e. the heterogeneous character of ADHD, has been pivotal in shaping the current theoretical understanding of the disorder. As first postulated by Sonuga-Barke (2002), in order to fit existing evidence theoretical models of ADHD must be able to fully account for the distinct and dissociable neuropsychological profiles associated with the disorder. More specifically, they must be able to explain how a multitude of genetic and environmental risk factors interact in order to give rise to the distinct patterns of executive dysfunction, aversion to delay, variable responses and timing deficits neuropsychological research has described. To accommodate these challenges, Sonuga-Barke (2002, 2003) initially formulated the dual pathway model of ADHD which described executive and motivational dysfunction as two distinct pathophysiological pathways of ADHD with different neurobiological bases, each independently contributing to the same final manifestation of the disorder.

Several other multi-deficit accounts of ADHD have been proposed since then by the same and by different authors (Coghill, et al., 2014; Fair, et al., 2012; Nigg & Casey, 2005; Sonuga-Barke, 2005). The original dual-pathway model has been extended to include a third timing deficit pathway (Sonuga-Barke, Bitsakou, et al., 2010) and to allow for more complex, synergistic interactions between the single pathways (Sonuga-Barke, Sergeant, Nigg, & Willcutt, 2008). This pattern of complexity is captured by the generic multiple-pathway framework described by Taylor and Sonuga-Barke (2008)(Figure 1.1). The emphasis here is on the complexity of ADHD pathophysiology: subtypes of neurocognitive dysfunction are seen as originating in the interaction between genetic and environmental risk factors and as being constantly mediated by additional environmental influences.

Figure 1.1. A generic multiple-pathway framework of the complex pathophysiology of ADHD *(From Taylor & Sonuga-Barke, 2008)*



The basic assumptions behind the dual-pathway model were empirically tested by Solanto and colleagues (Solanto et al., 2001) who demonstrated that executive functioning and aversion to delay were largely uncorrelated with each other but they were both associated with ADHD. Two neuropsychological measures were administered to a sub-sample of children participating in the large longitudinal Multimodal Treatment of ADHD (MTA) study (MTA Cooperative Group, 1999). The stop signal task was administered to 77 children to measure response inhibition and a choice delay task asking children to choose between a small immediate reward (1 point, delivered after 2 seconds) and a larger delayed reward (2 points delivered after

30 seconds) was used as an index of delay aversion. Discriminant function analyses indicated that the two tasks displayed modest discriminant validity when used separately (the rate of correct classification of children as “control” or “ADHD” was 61% for the stop-signal task and 53% for the choice delay task. But when both tasks were entered into the analysis in combination, the classification rates improved markedly reaching 88%. These findings have been widely interpreted as evidence for the independent association of EF and DA and have since been replicated in samples of preschool children (Pauli-Pott, Dalir, Mingebach, Roller, & Becker, 2013; Sonuga-Barke, Dalen, & Remington, 2003; Thorell & Wåhlstedt, 2006).

The ultimate aim of this modelling of heterogeneity is to aid diagnosis and improve treatment targeting (Taylor & Sonuga-Barke, 2008). By dissecting the clinical and neuropsychological heterogeneity of the disorder, these multifactorial models seek to identify groups of individuals who exhibit similar neurocognitive profiles, i.e. represent different “pathway types”. The underlying assumption is that individuals within these more homogeneous groups will be more likely to respond to treatment in similar ways, thus enabling the development of more targeted interventions, tailored to specific patterns of neuropsychological dysfunction.

The following chapter presents the dominant treatment model of ADHD. It places emphasis on the need for developing innovative, theory driven approaches that fit the current multifaceted conceptualisation of the disorder. It also makes the case that preventive early interventions represent a very good platform for developing exactly this type of treatment approach.

Chapter 2: ADHD treatments and the case for early intervention

2.1. Aim of the chapter

The aim of this chapter is to present the dominant treatment model for ADHD and its limitations and to lay out the theoretical rationale for early and preventive intervention as a new direction to ADHD treatment. In addition, it reviews the emerging evidence base for early intervention in ADHD.

2.2. The current treatment model for ADHD and its limitations

Evidence-based treatments for ADHD include medication, behavioural therapy and a combination of the two (Pelham & Fabiano, 2008; Taylor & Sonuga-Barke, 2008). Medication with stimulants and other pharmacological agents such as atomoxetine and clonidine has been extensively trialled and systematically reviewed (Banaschewski et al., 2006; Vitiello et al., 2001) and has been found to have large effects in reducing core ADHD symptoms for the majority of the children for whom it is prescribed in comparisons with placebo. However, there are several important limitations to it – each of which affect some children. First of all, treatment with stimulants is not helpful for many children either because they do not show a clearly beneficial response to it or because they experience side effects that do not allow the continuation of treatment, most commonly insomnia and appetite suppression (Graham et al., 2011). There is also the related issue of the unknown potential side effects of prescribing stimulants to young children for long periods of time, which is a particular concern for a substantial minority of parents (Sonuga-Barke, et al., 2006). Most crucially, longer term treatment effects lasting beyond the discontinuation of treatment have been difficult to demonstrate for the majority of children (Jensen et al., 2007) and there is little evidence to suggest that medication has beneficial effects on other areas of impairment, such as social relations and academic achievement (Pelham & Fabiano, 2008).

Behavioural therapy exploits the principles of positive and negative reinforcement to reduce ADHD-related difficulties and increase more desirable behaviours. It employs techniques such as paying positive attention to appropriate behaviours, setting negative consequences for problem behaviours and establishing effective communication rules. Behavioural interventions for the management of ADHD can be delivered in the family, as part of a parent training (PT) programme or in a school setting with teachers being instructed in the use of said techniques. One of the obvious advantages of psychological approaches is that they carry very little physical risk. They can also be “tailored” to match children’s particular treatment needs in specific domains of functioning and thus target broader areas of impairment –this is at least in theory if not always in practice.

Various reviews suggest that behavioural treatment does result in improvements of parent and teacher ratings of ADHD symptoms and of observed ADHD-related behaviours, with effect sizes varying from average to large (Antshel & Barkley, 2008; Chronis, Jones, & Raggi, 2006; DuPaul, 2007). On the other hand, a recent meta-analysis of behavioural interventions which differentiated between effects of treatment reported by raters closest to the therapeutic setting and thus prone to expectation bias (e.g. parents in the case of PT programmes, teachers in the case of school-based interventions) and raters more likely to be blinded to treatment (e.g. teachers in the case of PT programmes, or parents in the case of school-based interventions) found weak evidence for the efficacy of behavioural interventions in mediating core ADHD symptoms when rated by probably blinded raters (Sonuga-Barke et al., 2013). On the basis of existing evidence, and despite the fact that behavioural therapy is generally expected to be less effective in terms of core symptom reduction when compared with careful medication management (also see MTA Cooperative Group, 1999), current clinical practice guidelines recommend psychological treatment as the first choice in the treatment of ADHD, implemented alone or as a complement to medication (American Academy of Pediatrics, 2011; Taylor et al., 2004).

Among behavioural interventions, parent behaviour training has the strongest evidence base showing improvements in core symptoms in school aged (Antshel & Barkley, 2008; Chorpita et al., 2002) and preschool aged children (Charach et al., 2013; Knight, Rooney, & Chronis-Tuscano, 2008; LaForett, 2008; McGoey, Eckert, & DuPaul, 2002; Thompson et al., 2009), but also see Riddle et al. (2013) for a less promising outcome in a long term follow-up study of preschool ADHD. Still, compared to pharmacotherapy, parent training interventions still produce equal or larger effect sizes in preschool samples (Charach, et al., 2013). On the basis of this evidence and amid concerns regarding the safety and the long term efficacy of pharmacological intervention for preschoolers at risk for ADHD (Gleason et al., 2007; Kratochvil, Greenhill, March, Burke, & Vaughan, 2004; Vaughan & Kratochvil, 2006), there is strong clinical consensus that parent behaviour training should be the recommended treatment option offered to parents of very young children presenting with ADHD-related problems (American Academy of Pediatrics, 2011; Daley, et al., 2009; Ghuman, 2008; Kollins et al., 2006).

Behavioural approaches have important limitations too. Most notably, the majority of children who undergo behavioural therapy will not demonstrate maintenance or generalisation of treatment gains after the termination of treatment, as is the case with pharmacological treatment (Pelham & Fabiano, 2008). Another challenge is the effective continuation of treatment given that these approaches rely on the consistent use of behaviour modification techniques by parents and teachers, and of course the willing participation of the children themselves. This is an important barrier to optimal effectiveness, as evidenced by the existence of a “dose effect” where the greater the number of sessions attended, the greater the benefits of the intervention (Charach, et al., 2013). Finally, the cost of implementing behavioural therapy over long periods of time often makes such interventions relatively difficult to access (Daly, Creed, Xanthopoulos, & Brown, 2007; Koerting et al., 2013).

More recently, cognitive training has been put forward as an alternative approach to the management of ADHD symptoms. Cognitive training promises longer lasting changes to neurocognitive function by aiming to directly remediate specific cognitive deficits commonly associated with ADHD, such as deficits in attentional control, working memory and inhibitory control (Markomichali, Donnelly, & Sonuga Barke, 2009). Training typically involves extensive practice on a number of carefully selected training tasks, most frequently delivered by computers, over a period of time varying from a few days to several weeks. Crucially, these approaches include an adaptive element in that the difficulty of the training tasks is advanced gradually and in accordance with individual performance in order to make increasingly greater demands and thus ensure children are required to work at the limit of their ability at all times. The presumption therefore is that, through training, the underlying neural processes that are deficient in inattentive and impulsive children will be strengthened and in turn lead to generalised and long-lasting improvements in cognitive functioning (Klingberg, 2010).

The effectiveness of cognitive training, and of working memory training in particular, at enhancing cognitive capacity is currently a matter of debate (Melby-Lervag & Hulme, 2013; Shinaver, Entwistle, & Soderqvist, 2014; Shipstead, Redick, & Engle, 2012). Training related improvements in working memory function are fairly easy to demonstrate; the question is whether these improvements can have a longer lasting and further reaching impact on broader cognitive abilities. The effect of cognitive training approaches on ADHD symptoms was assessed by Sonuga-Barke and colleagues (2013). As was the case with behavioural interventions, when the analysis was limited to probably blinded assessments, the effects of cognitive training on ADHD symptomatology were very modest (Sonuga-Barke, et al., 2013). A more recent meta-analysis by the same group focusing not only on clinical but also on neuropsychological outcomes of cognitive training for children with ADHD reached similar conclusions. More specifically, it too confirmed that working memory training improves working memory function; but evidence for its effectiveness at reducing ADHD symptoms was again limited to unblinded assessments (Cortese et al., in preparation).

2.3. The case for early intervention in ADHD

In light of the limitations of frontline treatment options for ADHD, research into novel therapeutic approaches is still essential. Many researchers (eg. Nigg & Casey, 2005; Sonuga-Barke & Halperin, 2010) have advocated a translational approach to treatment development (Cicchetti & Toth, 2006; Curry, 2008) and have argued that the search for new treatments should be driven by the newly accumulated knowledge about the underlying pathophysiology of the disorder. Yet, it is a known and interesting fact that treatments for ADHD have very rarely, if at all, emerged as a result of a better understanding of the etiology of the disorder. Rather, most existing therapeutic approaches have been adapted from generic models of intervention borrowed from other clinical domains. Even recently developed interventions targeting very specific deficits associated with ADHD are still based on models or concepts developed outside the field of ADHD. One such example is attention training, which has been largely informed by generic models of attention as opposed to ADHD specific models (Sohlberg & Mateer, 2001).

The failure to link basic research with new treatment development in ADHD is clearly problematic. Sonuga-Barke and Halperin (2010) have made two propositions as to how to address this challenge. The first proposition is to seek new treatments within a developmental psychopathology perspective. Having identified the old bio-medical model as a barrier to translational science, they go on to argue that ADHD should be seen as a dynamic, complex outcome caused by the interplay of genetic and environmental risk factors and mediated by specific pathophysiological mechanisms and environmental influences (Sonuga-Barke, 2013). Understanding the diverse developmental risk-disorder pathways, such as the executive dysfunction and the motivational dysfunction pathway, is therefore considered to be an essential step in the development of new treatments as it could help predict the emergence of pathway-specific patterns of impairment, their mediating influences and ultimately their differential response to treatment (Willcutt, et al., 2008).

The second proposition is to focus on early intervention and prevention. A developmental psychopathology framework makes a useful distinction between causal processes (developmental risks) and outcome states (developmental outcomes, i.e. the disorder). Great emphasis is therefore placed on disorder precursor states and processes of alteration. These are seen as major intervention targets under the assumption that intervening early will be more effective than waiting until the pathogenic process is complete and the emergence of outcomes more established. Central to this notion is the hypothesis that environmental influences can facilitate and alter structural and functional brain development, which can in turn mitigate the expression of ADHD symptomatology over the course of development (Halperin, Bedard, & Curchack-Lichtin, 2012). Another important and related assumption is that if early intervention is successful and manages to alter the predicted developmental trajectories, this change should then be expected to be long lasting, a claim which has also been made in the context of other childhood disorders (Miklowitz & Cicchetti, 2006).

Both propositions are of particular relevance to this thesis. Under the first proposition, different (pathway-) types of impairment are expected to respond differentially to treatment. Children presenting with EF problems, for instance, are hypothesised to respond optimally to interventions targeting executive dysfunction. Indeed this seems to be the rationale for the development of tailored, process-specific executive function-training interventions for ADHD, such as working memory training, which have shown promising efficacy (Klingberg, 2010; Klingberg et al., 2005).

2.4. The evidence base for early intervention in ADHD

The first line of evidence in support of early intervention in ADHD comes from studies showing the benefits of preschool parent training. As discussed above, parent training interventions are already an established front-line treatment for preschool ADHD. The crucial question from an early intervention perspective is whether these interventions yield benefits lasting beyond the termination of active treatment. This seems to be a moot point at the moment with a recent long term follow up study (the Preschool ADHD Treatment Study, or PATS) reporting that the majority of 304 children assessed to be in the moderate-to-severe clinical range when they were, on average, 4.4. years old still met diagnostic and impairment criteria for ADHD at a three-, four- and six-year follow ups (Riddle, et al., 2013). Treatment in this study consisted of parent training followed by a double blind placebo-controlled medication phase. The absence of positive long term treatment effects serves as a reminder that preschool ADHD is a stable, chronic diagnosis and that more effective intervention strategies are needed for this age group. Still, parent training programmes, such as the Incredible Years (IY; Webster-Stratton & Reid, 2010) and the New Forest Parenting Package (NFPP; Thompson, et al., 2009) (a specialised, ADHD-specific programme which focuses on improving the quality of mother-child interactions) have in the past produced and maintained positive behavioural improvements in young children for periods of time up to 18 months after baseline (Jones, Daley, Hutchings, Bywater, & Eames, 2008; Sonuga-Barke, Daley, Thompson, Laver-Bradbury, & Weeks, 2001).

The NFPP encourages the management of ADHD symptoms through the use of parenting techniques that include joint play, turn taking and scaffolding and prescribes a number of games that parents and children can play together. Following similar principles, a number of innovative early interventions for ADHD are currently being tested which use games and game-like activities delivered by parents at home in order to promote positive parent-child interactions through joint play and to enhance neurocognitive skills that are frequently impaired in the disorder. In keeping with the view of ADHD as an executive dysfunction disorder, the great majority of these interventions have focused on executive function processes, such as working memory and inhibitory control. One such programme of activities is TEAMS (Training Executive, Attention and Motor Skills; Halperin et al., 2013), which is specifically designed for 4-5 year old children with ADHD and their parents. The programme teaches both parents and children games that aim to promote inhibition, attention, memory, planning and motor skills over a period of five weeks focusing on techniques that help parents tailor the programme’s activities to match the ability and skill of their children. An open pilot study of TEAMS has recently reported significant reduction in parent and teacher rated ADHD symptoms of inattention and impulsivity/hyperactivity, which persisted at a 3 month follow up (Halperin, et al., 2013). Parents’ satisfaction with the programme was also found to be excellent.

A similar programme, Enhancing Neurocognitive Growth with the Aid of Games and Exercise (ENGAGE; Healey & Halperin, 2014) is targeting three aspects of self-regulation that are known to be deficient in children with ADHD: behavioural, attentional and emotional control. What is of particular interest is that ENGAGE also incorporates activities that require physical exertion (e.g. skipping and ball games), alongside activities that require focused attention (e.g. working on puzzles) and behavioural control (e.g. playing “musical statues”). Again, all these activities are practiced by parents and children on a daily basis at home. The programme lasts for a period of five weeks. A preliminary study on the effectiveness of the ENGAGE programme has reported excellent compliance with the programme and improvements in parent-rated hyperactivity and in neurocognitive function, particularly in working memory and sensorimotor control. These improvements were maintained at a 12 month follow up (Healey & Halperin, 2014).

Further studies have examined the effects that child training alone (without parental intervention) has on early neurocognitive functioning. Much of this research is part of a wider developmental literature which has looked into the question of whether executive function skills can be influenced before children start school. These efforts are driven primarily by evidence suggesting that executive functions undergo a crucial and pronounced period of development between the ages of four and six years (Diamond, 2006; Rueda et al., 2004), that they are strongly associated with school readiness (Blair & Peters Razza, 2007; Neuenschwander, Roethlisberger, Cimeli, & Roebers, 2012; Ursache, Blair, & Raver, 2012) and that they reliably predict academic achievement in later years (Cartwright, 2012; McClelland & Cameron, 2011; Willoughby, Kupersmidt, & Voegler-Lee, 2012). Programmes aiming to promote EF development have been implemented in clinical, subclinical or typical samples of preschoolers and can be classified either as individualised or group training approaches (see Roethlisberger, Neuenschwander, Cimeli, Michel, & Roebers, 2012 for this classification), with group training being mainly curriculum interventions administered in educational settings.

In terms of individualised child training, there are only a handful of studies that have attempted to train EFs in young children and their outcomes are mixed. Dowsett and Livesey (2000) have demonstrated that three year old children with poor inhibitory skills can improve their performance on an untrained inhibitory task after brief practice in two EF tasks (two set shifting tasks). Similar practice related improvements were reported by Kloo and Perner (2003) in a sample of three year old children selected for having failed either of two criterion tasks: a card sorting task or a false belief task. Improvements were noted after children practiced on variants of the criterion tasks while receiving positive or negative feedback based on their performance. These findings have been recently replicated and extended by Espinet and colleagues (Espinet, Anderson, & Zelazo, 2013) who reported reductions in the amplitude of the N2 component of the ERP, an index associated with conflict detection, after a brief training session on the Kloo and Perner training protocol.

Other studies have used computerised training programs to promote EF development in young children with more limited success. Rueda and colleagues (Rueda, Rothbart, McCandliss, Saccomanno, & Posner, 2005) were not able to demonstrate any improvements in inhibitory control (measured by an untrained flanker task) in a sample of typically developing four-year-olds after five days of training. Training was on a computerised battery of executive attention tasks, which included stimulus discrimination, conflict resolution and interference control tasks. Improvements were only found on an IQ test (the Kaufman Brief Intelligence Test). Similarly, a study that investigated the differential effects of two types of EF training, WM and inhibition training, in a normative sample of preschoolers did not find any evidence of training effects (Thorell, Lindqvist, Nutley, Bohlin, & Klingberg, 2009). Despite improvements in inhibition tasks during training, neither of the two training programs that were tested led to any effects on two unpracticed inhibitory tasks, a stroop-like and a go/no-go task. But children who trained on the WM program did show improvements on trained and untrained WM tasks.

To date the strongest evidence that child training activities can improve children's EFs exist for two academic curricula: Promoting Alternative Thinking Strategies (PATHS; (PATHS; Riggs, Greenberg, Kusche, & Pentz, 2006)) and the Chicago School Readiness Project (CSRP; Raver et al., 2008). The PATHS intervention targets cooperation, emotional awareness, communication skills, self-regulation, self-esteem and problem solving in preschool children. The CSRP programme involves teacher training in behaviour management strategies designed to support children's self-regulation (such as implementing clearer goals and routines, rewarding positive behaviour, redirecting negative behaviour, etc.). Studies testing the value of these interventions have used robust methodologies (random assignment to conditions, active control groups and pre- and post-intervention assessment measures) and demonstrated convincingly that effects of training and practice generalised to objective measures of EFs on which the children had not practiced.

Less rigorous studies have also reported benefits to children's EFs from physical exercise (Chang, Tsai, Chen, & Hung, 2013), mindfulness training (Zelazo & Lyons, 2012), small-group play-based interventions (Roethlisberger, et al., 2012) and other early years academic curricula, such as the Tools of the Mind curriculum, which focuses on social, emotional and behavioural self-regulation (Diamond, Barnett, Thomas, & Munro, 2007), and the Montessori schooling system (Lillard & Else-Quest, 2006).

Taken together, findings from studies that seek to improve the executive functioning of typically developing preschoolers for the most part suggest that children's abilities to inhibit their actions, resolve conflict and keep things in working memory can be improved but the transfer of these skills to non-practised tasks has not always been possible to demonstrate. This could be because training was too brief, the training tasks were not age-appropriate or because the children were too young. Still, there are reasons for cautious optimism. Early indications suggest that the children who showed the most benefits from EF training were those with the poorest performances to begin with (see also Diamond & Lee, 2011). This finding makes it reasonable to assume that children with disordered inhibition, such as those at risk for ADHD, will benefit more from this type of training. Indeed, the preliminary results from play based interventions targeting preschoolers at risk for ADHD seem to be supporting this assumption.

The identification of early intervention targets in ADHD is still in its very early stages. The proposition that different causal pathways to ADHD underpin differential developmental trajectories to the emergence of the disorder has received empirical support, but to-date has had a minimal impact on therapeutic innovation. For instance, studies have shown that delay aversion and inhibitory control make independent contributions to preschool ADHD symptoms (Sonuga-Barke, Dalen, et al., 2003) and developmental outcomes (Thorell, 2007). But we are aware of very few attempts to develop tailored interventions and the methods that would allow us to reliably identify "children at risk for ADHD" as early as possible and to match them with the specific intervention they need, given the range of developmental phenotypes and pathways to the disorder.

In response to the need for more tailored early intervention approaches to ADHD and acknowledging the role of motivational dysfunction in ADHD, this thesis seeks to extend the scope of early intervention in ADHD by developing a training program that specifically targets sub-optimal motivational processes in young children at risk for ADHD. Children with altered motivational preferences are expected to benefit more from an intervention specifically designed to target motivational processes and this is exactly what the current thesis will attempt to demonstrate. The optimal time to implement this type of targeted intervention is early childhood. A number of research groups have independently developed interventions intended to promote neurocognitive functioning in very young children assessed to be at risk for ADHD. This emerging literature provides the evidence base for early intervention in ADHD and will be reviewed below. Following the same rationale, the intervention developed in the current thesis will also focus on children of a preschool age.

Chapter 3: Motivational dysfunction in ADHD as an early intervention target

3.1. Aim of the chapter

The aim of this chapter is to discuss the motivational dysfunction pathway to ADHD in more detail and to present an intervention model intended to isolate those aspects of motivational dysfunction that are most likely to be amenable to practice. To support the case for delay ability being a good target for early intervention, this chapter will review and synthesise evidence from diverse areas of psychological research suggesting that the motivational preferences of population groups prone to motivational failures, such as young children and adults diagnosed with addictive or disruptive behaviour disorders, can be influenced.

3.2. Motivational dysfunction in ADHD

Some of the earliest theoretical formulations of ADHD (e.g. Barkley, 1997a; Douglas & Parry, 1983, 1994) recognised an aberrant sensitivity to reinforcement as a key manifestation of the disorder hypothesising a reduced response to punishment and nonreward (Quay, 1997), an elevated reward threshold (Haenlein & Caul, 1987) and a sensitivity to removal of reward (Douglas & Parry, 1994). The resulting literature on effects of reinforcement contingencies on behaviour is small and rather fragmented and as such has produced largely contradictory evidence, with one notable exception: studies focusing on reward choice have produced convincing and replicable evidence showing that children with ADHD tend to choose smaller immediate rewards (smaller, sooner rewards; SS) over larger but more delayed rewards (larger, later rewards; LL) more frequently compared to typically developing children (Bitsakou, et al., 2009; Dalen, et al., 2004; Sagvolden, et al., 2005; Sonuga-Barke, Taylor, Sembi, et al., 1992; Tripp & Alsop, 2001), but also see Scheres et al. (2006) for an exception. Overall, the effect size for ADHD case-control differences has been comparable to those seen for executive function measures, making the preference for reward immediacy one of the most robust neuropsychological markers for ADHD (Willcutt, et al., 2008).

Intriguingly, neuroimaging data have revealed abnormalities in reward processing in ADHD. There is now a growing number of studies pointing to hypoactivation (i.e.

reduced activation) in the ventral striatum during reward anticipation in children and adults with ADHD compared to controls (Plichta et al., 2009; Scheres, Milham, Knutson, & Castellanos, 2007; Stroehle et al., 2008). Plichta and colleagues (2009) have also observed increased striatal and amygdala activations during delayed reward choices, marking a dissociation in striatal activation between immediate and delayed rewards. Further functional abnormalities associated with performance on delay tasks have been observed in the orbitofrontal cortex (OFC) and other limbic regions, such as the amygdala (Cubillo, Halari, Smith, Taylor, & Rubia, 2012; Paloyelis, Mehta, Faraone, Asherson, & Kuntsi, 2012). Whereas this body of evidence is not very large, increased knowledge about the reward circuitry of the brain is widely seen as supporting the notion that reward and delay related motivational processes constitute a significant and measureable deficit in ADHD (Kelly, Sonuga-Barke, Scheres, & Castellanos, 2007). It should not be a surprise, therefore, that a number of theoretical models have since sought to explain this motivational dysfunction pathway to the disorder, focusing in particular on ADHD children's preference for reward immediacy.

One dominant view attributes preference for reward immediacy in ADHD to disturbances in reinforcement learning and in particular impairments in the neural signalling of delayed rewards (Sagvolden, et al., 2005). These are hypothesised to result in a shorter and steeper delay-of-reinforcement gradient to that of typical children. The delay gradient refers to the function of reinforcement by time. A reinforcer is more potent when the time interval between the reinforcer and the response is short. A shorter and steeper gradient means that reinforcers need to be delivered very soon after a response is made in order to be effective in children with ADHD. Sagvolden and colleagues have argued that this is because the dopamine signalling system is weakened in ADHD and its ability to signal delayed future rewards is stunted (Sagvolden, et al., 2005).

Under a different formulation, ADHD children's choice preferences have been attributed to a different type of delay sensitivity. According to the delay aversion hypothesis, what motivates children with ADHD is not so much their preference for immediacy, but rather an aversion towards delay (Sonuga-Barke, Taylor, Sembi, et al., 1992). Models focusing on state regulation deficits in ADHD can also explain preference for immediate rewards. Even though state regulation deficits are typically defined in terms of imposed event rate conditions, choice of SS over LL rewards can be understood as an attempt on behalf of ADHD children to regulate their energetic state by choosing the delay setting that incurs an optimal activation level for them (Sonuga-Barke, Wiersema, van der Meere, & Roeyers, 2010).

Further to the above ADHD-specific models, it is also possible to use broader theoretical accounts of impulsivity and self-control to understand ADHD children's reward choices. For example, it is easy to see preference for reward immediacy as a function of discounting the value of future hypothetical rewards, in the context of the temporal discounting paradigm (Dias et al., 2013; Jimura, Chushak, & Braver, 2013), although delay discounting could also be linked to deficiencies in reinforcement learning. Indeed, tasks measuring the degree of delay discounting are reliable across time, sensitive and significantly correlate with impulse control conditions, such as obesity and ADHD (see for example Shiels et al., 2009). For these reasons, they are commonly used in investigations of the neural correlates of reward processing in adults with ADHD (e.g. Plichta, et al., 2009) and recent attempts have been made to adapt them so that they are suitable for use with children (Demurie, Roeyers, Baeyens, & Sonuga-Barke, 2012; Wilson, Mitchell, Musser, Schmitt, & Nigg, 2011).

Another related theoretical concept is that of the delay of gratification, which refers to the ability of young children to postpone immediately available gratification in order to pursue more attractive later rewards in a classic laboratory task popularly known as the "marshmallow test". This task simply measures the amount of time a child can resist opting for a small, immediately available reward (one marshmallow) in order to obtain a larger, delayed reward (two marshmallows, available at the end of the waiting period, some 15 minutes later). From a personality theory perspective, this ability is seen as necessary for the prevention of self-regulatory failures often presumed to be associated with antisocial behaviours, conduct disorders, eating disorders, drug abuse, lack of resilience and others (Mischel, Shoda, & Rodriguez, 1989; Peake, Hebl, & Mischel, 2002). Within these broader theoretical frameworks ADHD children's preference for immediate rewards is understood as a breakdown of self-control and it is not uncommon for ADHD to be discussed in conjunction with other "impulse-control" disorders both in terms of our theoretical understanding of it and in terms of preferred treatment options (Dichter, Damiano, & Allen, 2012; Strayhorn, 2002a).

The fact that many of these theoretical accounts are in agreement over some of their most basic predictions makes it difficult to discriminate between them. In recognition of this fact, there have been calls for more specific and experimentally testable theoretical predictions to help guide future research in the field (Luman, Tripp, & Scheres, 2010). An additional complicating factor is that many of these models operate at very different levels of explanation, with some focusing on behavioural outcomes, others on the underlying neurobiology of reward and very few attempting to provide multilevel explanations spanning intermediate levels of investigation as well. Multimodal research methods is evidently the direction the field is moving toward, and there are already promising pieces of research investigating reward-related processing in ADHD using combined behavioural and neuroimaging methods in order to clarify old theoretical notions (see for example Wilbertz et al., 2013) or test clear theoretical predictions (van Meel, Heslenfeld, Oosterlaan, Luman, & Sergeant, 2011).

Further disambiguation of these theoretical accounts is beyond the scope of this thesis. Rather, its particular focus is to isolate those aspects of motivational dysfunction in ADHD that have been shown to be more amenable to practice and could thus serve as potential intervention targets. In the remainder of this chapter, impulsive choice will be looked at from different perspectives in order to review and synthesise evidence from diverse areas of psychological research suggesting the motivational preferences can be influenced through practice.

3.3. Influencing impulsive choice

Perhaps not surprisingly, the question of how to teach children to wait longer in exchange for longer term goals spans different fields of psychological research and many decades. B.F. Skinner, for example, described exercises of teaching children to wait (when tired and hungry above a bowl of steaming soup) as part of his psychological utopia in Walden Two (Skinner, 1948). Self-control failures are commonly seen as the central deficit of many impulse control disorders. From a clinical perspective, therefore, a great number of behavioural therapy interventions for disruptive behaviour disorders of childhood could be seen as self-control training programs (Strayhorn, 2002b). But not all such interventions have focused on behavioural therapy alone. Many other research streams have explored alternative, training-based strategies for improving children's delay tolerance and these will be the focus of the review that follows.

3.3.1. Impulsivity and the shaping and fading of delay

Organisms are seen as able to self-control when, in a choice paradigm, they prefer larger, more delayed rewards over smaller and more immediate ones. The antithesis of this skill (i.e. preference for immediate rewards) has been defined as impulsivity. Importantly, the key factor hypothesised to determine reward preference is the delay occurring between choice and reward, with preference for LL rewards diminishing as the delay increases. In this context, a motivational style that consistently favours immediacy over delay, characteristic of ADHD and other impulse control conditions, can been conceptualised as a deficit in inhibitory control (i.e. difficulty withholding the SS response; Barkley, 1997a) or as a steep function relating reward magnitude and delay, with rewards becoming less effective the more delayed they are (Sagvolden, Aase, Zeiner, & Berger, 1998).

Choice paradigms of self-control have been utilised to explore questions as to which species under which conditions are capable of mastering this skill. Self-control training procedures typically involve long series of alternative choices between reinforcers (most often consumable rewards) differing in size of reward, length of delay and type of delay (e.g. pre- or post-reward). The rate of reinforcement delivery can also be manipulated and be constant or probabilistic (Mazur, 2005). Logue (1995) has described experimental procedures which result in pigeons learning how to delay gratification. One procedure involves rewarding very gradual changes toward the desired behaviour. For example, pigeons are initially given the choice between two rewards, one of which is better than the other (tastier or larger). When the bird is accustomed to choosing the better reward, a delay interval is introduced between the choice and the reward, which is subsequently gradually increased (delay shaping). It is also possible to start off by letting the pigeon choose between two delayed rewards, and then gradually make the less preferred reward more immediate (delay fading).

Interestingly, similar reinforcement schedules have in the past been used to encourage a preference for delayed rewards in children. The idea here is the same: children can learn to opt for delayed rewards after being gradually exposed to longer and longer delays. In a study by Schweitzer and Sluzer-Azaroff (1988) five impulsive children (as identified by their teachers) aged four to six years were taught to overcome their initial preferences for immediate rewards by being exposed to gradually increasing delays over many sessions. The study used an apparatus for the presentation and the delivery of the rewards, which did not allow children access to the rewards before the termination of the delay periods. Training was delivered in a school setting in 2-3 sessions per week and lasted for four months on average (until a termination criterion was reached or until the end of the school year). Children chose what rewards they wanted to work for at the beginning of each session (stickers or edible treats or a combination of the two). Delay values for the two reward choices (one vs three items) began at 0 seconds each, but the delay for the larger reward was increased by 5-second increments every time the child reached a success criterion. This was reached when the child selected the larger reward on four out of five trials. As expected, the delays that children were able to tolerate in order to get the larger reward increased (they ranged between 20 and 65 seconds at the end of training). But the most dramatic changes were recorded when children’s indifference points before and after training were compared (indifference points are delay values which result in the child choosing either reward option equally often). For three out of the five children indifference points were not acquired at the post-test because these children kept choosing the larger reinforcer even at delays of 60 seconds. For the remaining two children, indifference points increased from 1.7 to 37.5 seconds and from 25.9 to 47.1 seconds. Similar procedures have been used successfully since then to modify reward preferences in adults with developmental disabilities (Dixon et al., 1998; Dixon & Tibbetts, 2009) and to replicate the above findings in a series of case studies of children with ADHD (Binder, Dixon, & Ghezzi, 2000).

Neef and colleagues have demonstrated that a delay fading procedure can increase preference for delayed rewards in ADHD using a slightly different paradigm (Neef, Bicard, & Endo, 2001). In the pre-test phase of this case series study, the relative influence of reward delay was assessed in comparison with other competing reward dimensions such as reward quality, rate of reward delivery and task effort. This was done by asking children to choose between two math questions they would like to work on, each of which was accompanied by a different reward option. Every reward dimension was set against all the others, in pairs. Interestingly, results from this phase indicated that reward delay (delayed rewards were available only at the beginning of the next experimental session, a day later) was the most influential reward dimension for all three children (this was measured as the percentage of time each child allocated to each reward option). In the training phase, reward immediacy was initially set at 15 minutes and competed with the next most influential reward dimension for each child. Delay was subsequently increased in increments of 15 minutes when a success criterion was reached (70% of time allocation to this reward option over two experimental sessions). Children were reassessed after training to determine their preferences among all reward dimensions (in a replication of the pre-test conditions). Results indicated that children now favoured rate and quality of rewards over reward immediacy and low task effort even when rewards were delayed for as long as 24 hours.

Specifically with regard to ADHD, a lot of interest has focused on identifying contextual and environmental factors that may influence ADHD children's patterns of impulsive choice. For example, it has been shown that the availability of other rewards or other types of stimulation during delay periods can increase self-control in children with ADHD (Grosch & Neuringer, 1981; Schweitzer & Sulzer-Azaroff, 1995). The opposite was true for manipulations of the novelty of the experimental setting and the time children have spent on task. More specifically, decrements in performance of children with ADHD have been found as the novelty of a situation decreases or when repeated measures are taken (Alberts & van der Meere, 1992; Zentall & Zentall, 1976).

Taken together these findings suggest that children's ability to self-control can be strengthened by intensive practice. Of course, the applicability of such intensive, lengthy and monotonous reinforcement schedules in educational or clinical contexts can be challenging and further research is needed to identify how the principles of reinforcement training might be adapted to fit more realistic conditions and time scales. There is also the obvious question of whether the effects of intensive training are generalizable beyond the particular context in which learning has occurred, particularly given the importance of examining impulsive choice with the broader context of other reward contingencies.

3.3.2. Delay of gratification and the strategic deployment of attention

Children’s ability to self-control and delay gratification (DG) has been ingeniously operationalised in a slightly different manner. The classic marshmallow task also gives children a choice between two reward options (an SS and an LL one), but they only ever get one chance at choosing. More specifically, in the classic DG task an experimenter offers children a small immediate reward of an edible treat that can be eaten right away (e.g. one marshmallow), or a larger reward (e.g. two marshmallows) that can be obtained only if children can refrain from eating the small reward and wait for the experimenter to return to the room after she is gone for a few minutes. Both rewards are typically left in front of the children for the duration of the waiting period, but this is an aspect of the protocol that has often been the object of experimental manipulation. Children are also given a bell to ring to summon the experimenter to the room whenever they wish. The latency until the end of the waiting period (15 minutes in a recent protocol (Eigsti et al., 2006); up to 20 minutes in earlier studies (Moore, Mischel, & Zeiss, 1976)) or until the moment children ring the bell or start eating the snacks is the main index of DG.

Modified DG protocols have used alternative resistance-to-temptation scenarios (for example, children are shown an attractive toy but they are not allowed to play with it until the experimenter is back (in Carlson, 2005) and have added more trials to the classic DG task. For example, in the Cookie Delay task a small edible treat is placed in front of the child who is then asked to wait for a signal before she can retrieve it. In the standard version of this task, this situation is repeated eight times with delays varying between 5 and 30 seconds applied each time (Campbell, Szumowski, Ewing, Gluck, &

Breaux, 1982). It is easy to see that reward choice and DG tasks are measuring very similar constructs. Some of the standard DG tasks are perhaps more suited to social psychology research and to addressing questions of a relational nature on account of being essentially one-trial tasks, but the newest versions of these protocols are improved and are widely used to measure delay tolerance in normative and clinical samples of very young children.

Longitudinal studies into the development of children's ability to delay gratification have produced some impressive evidence, particularly in terms of the association between performance in the preschool task and educational and clinical outcomes in adolescence (Mischel, Shoda, & Peake, 1988; Mischel, et al., 1989). More specifically, the amount of time children were able to wait in the preschool task was found to correlate significantly with their Scholastic Aptitude Test (SAT) scores ten years later (Pearson's $r = 0.42$ for verbal scores and $r = 0.57$ for math scores). Children who did better at the preschool task were also rated to be more self-controlled and more able to cope with frustration by their parents compared to adolescents who did worse.

Studies conducted within the DG framework suggest there are two ways to help children sustain the delays for longer: the first is by diverting children's attention away from the rewards they are striving to achieve by waiting. Mischel and Ebbesen (1970), for example, have shown that exposure to rewards or cues to attend to rewards during delay undermine children's ability to wait successfully. In a study where children's thoughts were cued either to focus on the rewards or to distract them from the rewards, results showed that delay times were longer (around 10 minutes on average) when children were instructed to think fun thoughts compared to when they were instructed to think about the rewards (mean delay time was found to be less than 5 minutes in this condition). Similarly, distracting children from the rewards by giving them a fun toy to play with also helped them increase their waiting times (Mischel, Zeiss, & Ebbesen, 1972). The second method that has been used to facilitate delay of gratification is to use cognitive reframing to divert children's attention away from the motivating, hard-to-resist, "hot" qualities of the rewards. For example, when children were taught to turn the treats "into pictures in their heads" or to think about "how puffy marshmallows are, like cotton balls or clouds" their delay times improved dramatically compared to condition where they had been cued to think about “how gooey and yummy marshmallows taste”.

Collectively, these findings suggest that children are helped to sustain goal-directed delay when their attention is directed away from the rewards they are waiting for or when they are helped to adopt an abstract, “cool” focus on the rewards. The use of a hot-cool framework in which “cooling” and “heating” strategies are seen as necessary for maintaining self-regulation (cooling strategies are necessary for exerting self-control and heating strategies for constantly renewing commitment to the pursue of long term goals) has since been extensively used to guide clinical psychology research (see Kross, Ayduk, & Mischel, 2005 for an example). These are important findings that merit further investigation and have obvious clinical value, but from a training perspective they pose some interesting questions. Firstly, it is important to note that the majority of DG tasks are one trial protocols with uncertain test-retest reliability. The role of using explicit strategies to guide children during training is also a point worthy of careful consideration. Research has shown that self-distraction strategies will help children sustain longer delays by focusing their attention elsewhere.

3.3.3. The discounting of delayed rewards in clinical contexts

Temporal or delay discounting (DD) refers to the wider concept of self-control and deferred gratification. This notion refers to the decrease of the subjective value of a reward as a function of the time until its receipt. The degree at which individuals discount delayed rewards can be estimated with remarkable accuracy by experimental procedures requiring participants to respond to hypothetical reward choice scenarios, typically offering choices between two amounts of money, large or small, to be delivered in the near or distant future. Across numerous trials the amount of the small rewards and the time of delivery of the large rewards are manipulated so that the exact subjective value of the delayed rewards can be identified. This is done by determining the point of indifference between the magnitude of a small reward available immediately and the delay in the delivery of a larger reward, which is considered an indicator of the extent to which a large reward has been discounted because of its delayed delivery point (Critchfield & Kollins, 2001).

Because they allow reward choices to be precisely quantified and scaled in time, DD procedures are widely regarded as very sensitive measures of impulsivity and are extensively used in a variety of contexts ranging from clinical settings to the study of behavioural economics. From a personality theory perspective, the degree of delay discounting is relatively stable across time and contexts, so that some theorists see it as a personality trait (Odum, 2011). Odum defends this position by citing findings showing that: people who discount one type of reward (for example money) also tend to discount other types of rewards (for example food; Tsukayama & Duckworth, 2010); that the degree of discounting is similar across real and hypothetical rewards (Johnson & Bickel, 2002); that steep delay discounting is related to an array of maladaptive behaviours, such as substance abuse, gambling, obesity and others (see Yi, Mitchell, & Bickel, 2010 for a review) and that there may be a genetic basis to this characteristic with different strains of rats and pigeons showing differing degrees of discounting (Wilhelm & Mitchell, 2009) and high heritability of this trait in humans (Anokhin, Golosheykin, Grant, & Heath, 2011).

Finding ways to influence the degree of delay discounting could have broad clinical implications. As discussed above, impulsive decision making can be modified through shaping and fading procedures that manipulate reward magnitude and delay. Yet, a small number of recent studies have reported significant reductions in delay discounting by using no direct manipulation of the delay discounting procedure itself. Past research has demonstrated strong associations between working memory and degree of delay discounting (Shamosh et al., 2008), leading researchers to hypothesise that improving working memory function could result in less discounting of future rewards. One such study has shown that working memory training was successful at reducing degree of delay discounting in individuals who were receiving treatment for stimulant use (Bickel, Yi, Landes, Hill, & Baxter, 2011). Indeed, after training on a commercially available battery of working memory tasks (PSSCogReHab) for an average of 25 days, participants in the active control group significantly decreased their discounting rate k by 50% relative to the performance of an inactive control group. A second study used a money management training programme to demonstrate increases in the valuation of future rewards (i.e. less impulsive decision making) in a group of psychiatric patients with histories of substance abuse using a DD task (Black & Rosen, 2011). The reason this intervention was selected was that it had been proved effective in reducing levels of substance use in previous studies.

Despite the relative stability of degree of discounting further studies have indicated that DD can be influenced by numerous contextual and individual state factors, such as exposure to images of natural environments (as opposed to images of man-made environments or of geometrical shapes) (Berry, Sweeney, Morath, Odum, & Jordan, 2014) and levels of blood glucose, with participants showing reduced levels of discounting after consuming a sugared drink compared to a control group who had consumed a similar drink sweetened with artificial sweeteners (Wang & Dvorak, 2010). Discounting is also known to depend on framing manipulations, for example on whether preferences are elicited by binary choices (e.g. "would you rather have x amount of money today, or y amount of money in one month?") or open ended questions (e.g. "what is the minimum amount you would rather have in a year instead of x amount today?") (Frederick & Loewenstein, 2008) and on recency effects (Prelec & Loewenstein, 1991). Such findings have led some authors to emphasize how variable discounting is (see Frederick & Loewenstein, 2008 for example) and how sensitive to many known (and possibly even more unknown) contextual factors. From a training perspective, this observation is crucial as the effects of interventions aiming to influence degree of discounting are likely to be contingent on these contextual factors as well.

3.3.4. A note on self-regulation and the role of meta-cognition

Self-regulation is a broad and multidimensional construct which refers to both cognitive and behavioural processes through which individuals are able to maintain optimal levels of emotional, motivational and cognitive arousal (Blair & Diamond, 2008). It is a term used particularly frequently in the educational and developmental sciences, but it also has applications in the ADHD field. A recent review article, for example, positioned self-regulatory deficit models, such as the cognitive-energetic model (Sergeant, 2000), as an alternative to the core cognitive/motivational models of ADHD because they integrate cognitive and motivational processes in a single etiological model of ADHD (Shiels & Hawk, 2010). But despite an explosion of research interest into the emergence of self-regulatory skills in early childhood, the field has been hindered by a marked lack of conceptual clarity and poor operationalization (McClelland & Cameron, 2012; Schunk, 2008).

Of relevance to the current discussion are interventions designed to address self-regulation deficits in ADHD. These often include cognitive training procedures such as self-monitoring, self-reinforcement and self-management (Reid, Trout, & Schartz, 2005). Broadly defined, these interventions aim to increase awareness of behaviour by asking children to observe and record aspects of their levels of attention and or performance. Despite earlier studies concluding that the evidence in support of cognitive interventions in ADHD is lacking (Abikoff, 1991), more recent studies have yielded more promising results (Guderjahn, Gold, Stadler, & Gawrilow, 2013; Reid, et al., 2005). Many of these strategies do not target impulsive choice directly, but some of them do address issues like goal-pursuit and goal realisation and as such are often used to support individuals whose self-control competences are diminished in academic or clinical contexts.

3.3.5. Delay aversion and the delay exposure model for intervention

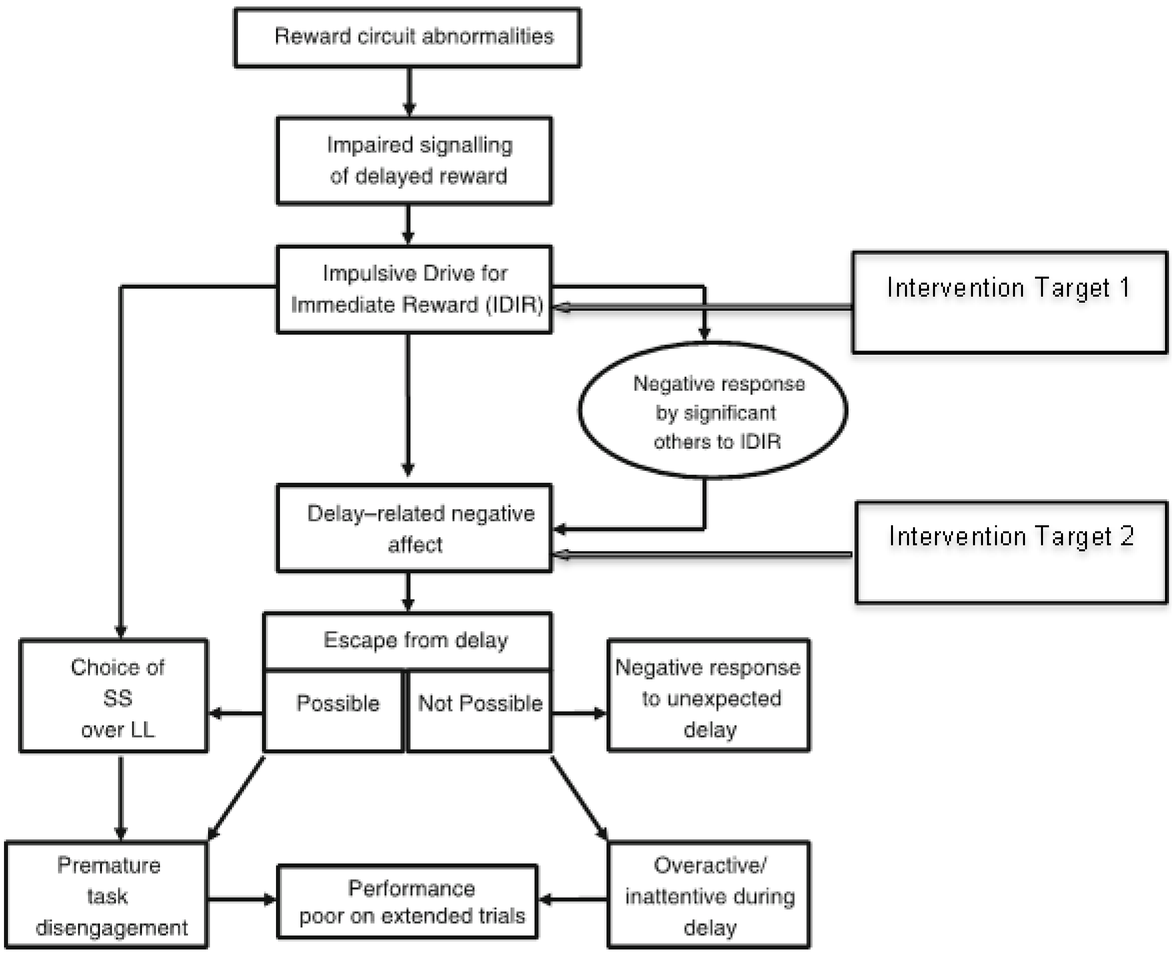
The delay aversion (DA) hypothesis was initially formulated to account for ADHD children's preference for immediate over delayed rewards (Sonuga-Barke, Taylor, & Heptinstall, 1992; Sonuga-Barke, Taylor, Sembi, et al., 1992). It postulated that the primary motivator of this pattern of behaviour in children with ADHD is not their impulse to choose the immediate option, but their aversion towards the delay associated with the delayed option. Put differently, according to the DA hypothesis, when ADHD children are making reward choices their goal is to minimise their time on task, in contrast to control children who are typically motivated by maximising their total gains.

The DA hypothesis can be elegantly tested using standard choice delay tasks. By imposing an obligatory period of post-reward delay after the delivery of the SS rewards, the SS option no longer reduces the overall amount of delay incurred by any particular reward choice. This manipulation allows comparisons to be made between reward preferences in conditions with or without post-reward delay to determine the relative influence of DA when escape from delay is possible and when it is not. According to the DA model, delay averse children would be predicted to choose the SS reward more often when it is possible to minimise delay (in no post-reward delay conditions) compared to when it is not (in post-reward delay conditions). This prediction was met with empirical support at the time (Sonuga-Barke, Taylor, Sembi, et al., 1992) and has been replicated numerous times (Marco et al., 2009; Tripp & Alsop, 2001).

Since then the delay aversion hypothesis has been developed into a broader developmental model of ADHD emphasising the existence of multiple etiological pathways to the emergence of the disorder and focusing in particular on the emergence of inattention and hyperactivity symptoms in non-escape delay situations (Sonuga-Barke, 2003, 2005). The motivational pathway distinguishes between primary manifestations and secondary adaptations of DA (Figure 3.1). Primary manifestations of DA are seen as driven by deep seated impairments in the dopamine-mediated reward circuitry of the brain that disturb the processing of delayed rewards and lead to systematic preferences for reward immediacy. This mechanism is closely related to the hypofunctioning dopamine systems model described by Sagvolden and colleagues (2005). The resulting pattern of reward choice behaviour has recently been described by the narrow term 'impulsive drive for immediate reward' (IDIR), in juxtaposition to the wider concept impulsiveness with its many different connotations (in Marco, et al., 2009). According to the DA model, this fundamental IDIR-related deficit is subsequently compounded by a generalised negative emotional response to delay rich environments or situations. Put simply, children with ADHD may begin life with particular motivational tendencies (e.g. to value delayed rewards less) but they become delay averse (i.e. learn to avoid delay) as they grow up because they come to associate delay rich situations with negative affect (Sonuga-Barke, 2003). Importantly, both IDIR and DA are seen as primary motivational processes that contribute to impulsive choice, i.e. they act together to create the marked preference for reward immediacy characteristic of ADHD.

Secondary adaptations of DA refer to behaviours expressed in non-laboratory, everyday life settings as a function of whether escape from delay is possible or whether delay must be tolerated (Sonuga-Barke, 2005; Sonuga-Barke, et al., 2008). When delay is avoidable, the model predicts that delay averse children will systematically avoid it in order to achieve a reduction in negative affect (by investing less time and effort in a task despite the associated costs of such behaviours). When delay has to be tolerated, they may engage in behaviours intended to reduce the subjective passage of time by directing their attention to non-temporal stimulation (Antrop et al., 2006) or they may become frustrated, more distractible or hyperactive, particularly when faced with unexpected and unavoidable delay (Bitsakou, Antrop, Wiersema, & Sonuga-Barke, 2006).

Figure 3.1. An integrated model of delay aversion
(Sonuga-Barke, Wiersema, van der Meere, & Roeyers, 2010; Sonuga-Barke, 2004)



3.3.5.1. Supporting evidence

The evidence in support of the DA model is extensive. Compared to typically developing children, children with ADHD show a marked preference for immediate over delayed rewards (Antrop, et al., 2006; Bitsakou, et al., 2009; Dalen, et al., 2004; Marco, et al., 2009; Schweitzer & Sulzer-Azaroff, 1995; Solanto, et al., 2001; Sonuga-Barke, Taylor, Sembi, et al., 1992; see also Willcutt, et al., 2008 for a review); they also become more frustrated by unexpected impositions of delay (Bitsakou, et al., 2006); show increased activity levels during delays (Antrop, Roeyers, Van Oost, & Buysse, 2000); are more vigilant to environmental delay related cues (Sonuga-Barke, De Houwer, De Ruiter, Ajzenstzen, & Holland, 2004); and their performance is affected by demanding tasks utilising slow event rates (Wiersema, van der Meere, Roeyers, Van Coster, & Baeyens, 2006).

Marco et al. (2009) investigated the issue of the relative contributions of DA and IDIR to ADHD children's preference for immediate rewards. The results demonstrated that children with ADHD chose SS over LL consistently across delay conditions (i.e. when pre- and post-reward delays where imposed) but did more so when delays were avoidable (i.e. when no post-reward delays were applied). These findings support the notion that both IDIR and DA are necessary to explain reward choice in ADHD. Very recently, a multimodal investigation used behavioural, physiological and neuroimaging data to demonstrate that children with ADHD display an exacerbated emotional state during the anticipation and experience of delays (Wilbertz, et al., 2013). There are still some aspects of the DA model that have not been subjected to experimental testing, with the developmental dimension of the model perhaps being the most significant one (i.e. the assertion that DA and its secondary adaptations should arise subsequently to the primary neurobiological IDIR-related deficits).

3.3.5.2. The delay exposure intervention model

According to the DA hypothesis there are two distinct stages in the emergence of impulsive choice in children with ADHD: an initial motivational tendency to favour immediate over delayed rewards, linked to deficits in the neural signalling of delayed rewards, and the subsequently developed negative affect associated with delay. By extension, Sonuga-Barke (2004) has suggested that interventions aiming to alter ADHD children's motivational style should address both elements of impulsive choice and therefore have two intervention targets, IDIR related abnormalities and delay related negative affect (see also Figure 3.1). To address the IDIR, he proposes that training programmes should aim to strengthen the association between a response and its delayed reward outcome. This association is hypothesised to be weak in children with ADHD and strengthening the link through delay exposure (i.e. extensive practice at receiving delayed rewards) could potentially have an impact on or even normalise the dopamine signalling system. To help children overcome their aversion to delay, exposure interventions should aim to expose children to repeated instances of achievable, rewarded delay. This is proposed in the hope that repeated presentations of aversive delay related events will eventually lead to habituation (a form of delay desensitization) and subsequent reduction in delay avoidant behaviours. Importantly, this intervention model implies that *delay exposure* is the active ingredient of this type of intervention, as opposed to strategies aimed at helping children cope with delay. Consequently, it advocates an implicit training approach as the most appropriate method for restructuring delay preferences and promoting delay tolerance in children presenting with an impulsive motivational style.

Amidst calls for improved, theory driven and neuropsychologically informed interventions for ADHD (Rutledge, van den Bos, McClure, & Schweitzer, 2012), Sonuga-Barke's proposition that a renewed focus on delay exposure training may constitute a therapeutic possibility for intervention in ADHD seems timely and appropriate. Its rationale is based in recent advances in ADHD theory and research and addresses the heterogeneous character of the disorder. Furthermore, it targets multiple elements of impulsive choice making it one of the most comprehensive approaches out of those aiming to modify motivational preferences in children. It is also in keeping with the main principle of early intervention seeking to develop interventions that target the diverse developmental pathways to ADHD.

For all these reasons, it is the expressed aim of the current thesis to develop an intervention to target the motivational pathway to ADHD in accordance to Sonuga-

Barke's delay exposure model. The process of developing a delay training paradigm in line with the principles and specifications of this model will be described in detail in Chapter Five. The intervening Chapter Four will present the psychometric properties of two new measures of delay tolerance developed in this thesis.

Chapter 4: Measure Development: two novel measures of delay tolerance in young children and their psychometric properties

4.1. Aims of the chapter

The aim of this chapter is to introduce two new measures of delay tolerance to be used with children of a preschool age and to examine and evaluate their psychometric properties in the context of the challenges associated with measuring delay tolerance and aversion in this age group. The first of these instruments is the Bee Delay Task, a computerised delay task designed specifically for this thesis in order to improve the sensitivity of choice delay tasks in this age. The second instrument is the parent/teacher version of the Quick Delay Questionnaire (QDQ), which was originally developed as a self-report questionnaire assessing delay related behaviour in adults. This was adapted into a rating scale measuring children's delay related difficulties designed to be completed by adult informants (parent, teachers or caregivers).

4.2. Measurement of delay aversion in young children

Given its central role in many models of psychological development of self-control it is not surprising that the child's response to delay has received considerable attention and has been measured in a number of different ways. As discussed in the previous chapter, these have included primarily the classic marshmallow task which measures the amount of time children are willing to wait in order to obtain a more valued but delayed outcome (Mischel & Metzner, 1962) and tasks which give participants a series of choices between small immediate rewards and larger rewards presented after a period of delay (typically lasting around 30 seconds)(see for example the Maudsley Index of Delay Aversion (MIDA) task in Kuntsi, Stevenson, Oosterlaan, & Sonuga-Barke, 2001). These choice delay tasks typically calculate the percentage of times the larger, delayed reward is selected as an index of impulsive choice.

More recently, these approaches have been complemented by experimental paradigms focusing on aspects of delay tolerance that go beyond the indexing of impulsive choice. The Delay Frustration Task (DeFT) measures the frustration participants feel when unexpected and unsignalled delays disrupt the completion of what they view as the primary task they have to complete (Bitsakou, et al., 2006). In this task, participants are instructed to answer a computerised set of simple multiple-choice math questions. On a minority of the trials the computer program appears to freeze and postpones access to the next question by 3 to 20sec. The density of response per second (determined as the product of the number of times participants pressed the button to move to the next question while the computer was inactive by the duration of each button press) is the outcome measure of this task. The Delay Reaction Time (DRT) task is another task that does not involve impulsive choice (Bitsakou, et al., 2009; Sonuga-Barke & Taylor, 1992). It measures the impact long presentation times (event rates) have on reaction times as a means of indexing delay aversion. The task requires participants to respond to pictorial stimuli as accurately and as quickly as possible. The presentation rate of the stimuli is manipulated, with some stimuli presented after considerable delays of up to 20 seconds. The main index of the DRT task is the participants reaction times obtained after delaying the stimulus presentation minus the typical reaction times on the no delay trials.

The psychometric properties of these tasks are largely satisfactory. Choice delay tasks can differentiate between ADHD cases and controls to a satisfactory degree, with pooled effect sizes ranging from 0.57 to 0.71, depending on the type of task used (Willcutt, et al., 2008). These effect sizes are comparable to those seen for executive function measures. Further to that, the test-retest reliability of the percentage of LL responses in the MIDA, of the response frequency in the DeFT and of RTs in the DRT have been found to be very high: interclass correlation = .74 (Kuntsi, et al., 2001), α coefficient = .92 (Bitsakou, et al., 2006) and intra-class correlation = .79 (Bitsakou, 2007), respectively.

Despite the stability of delay performance over time, further issues interfere with the validity and age sensitivity of delay tasks. For a task to have construct (or convergent) validity it needs to correlate with another criterion measure known to be valid from the literature (such as a validated questionnaire) or with pre-existing tasks measuring related constructs as specified by theory. With regard to this latter point, while choice of small over large delayed rewards is widely seen as the hallmark of delay aversion, other responses to the imposition of delay have also been predicted to reflect delay aversion, such as heightened emotional responses to delay-rich settings and attentional biases to delay related cues. Such effects have been postulated according to theory (Sonuga-Barke, 2003; Sonuga-Barke, Houlberg, & Hall, 1994) and have received empirical support (see for example Bitsakou, et al., 2006; Sonuga-Barke, et al., 2004). But they also contribute to the multifaceted nature of the construct and as a result make delay aversion difficult to measure as a single unitary construct across different ages. Indeed, studies that have examined the association between different delay tasks, have found either low correlations between them (Bitsakou, et al., 2009) or have revealed multiple components of delay challenging the single factor hypothesis (Sonuga-Barke, Bitsakou, et al., 2010).

To date the study of delay tolerance in ADHD has been largely focused to the school aged period. Attempts to use some of these procedures with children of different ages have been met with limited success. Multiple studies have reported ceiling effects in choice delay tasks, with adolescents having little difficulty choosing the LL response compared to younger children (Bitsakou, et al., 2009; Marco, et al., 2009). This could be either because the small monetary rewards typically used in choice delay tasks cease to be reinforcing at this age or because the ability to tolerate delays grows sharply in adolescence (Bjork et al., 2004). Of specific interest to this thesis is the existence of age effects at the other end of the age bracket in tasks used in the preschool years. Choice delay tasks involve a constellation of cognitive skills in addition to delaying gratification, including verbal comprehension, critical reasoning, decision making and response inhibition among others. Because all these skills develop rapidly during the preschool years, it is crucial that delay tasks are specifically assessed for use in early childhood. There is already anecdotal evidence suggesting that when children who are too young to grasp the delay to reward size trade-off are faced with a series of SS/LL choices, they tend to alternate their responses and perform at the level of chance (as opposed to favouring the SS option as one might have expected).

In response to these considerations, a novel choice delay task was designed specifically for this thesis with young children in mind with the primary objective of making delay choice less abstract and easier for young children to grasp by removing the need for them to choose between two alternatives. The current thesis also embarked on the development and validation of a teacher rating scale to enable the measurement of aspects of delay aversion other than those directly assessed through neuropsychological testing.

4.3. Development and initial validation of the Bee Delay Task

The Bee Delay task is an adjusting delay task. It is designed to measure children's delay tolerance by calculating the amount of time children are willing to wait for delayed rewards. It is also designed in such a way as to make delay choice, i.e. the trade-off between reward magnitude and delay, easier for children to understand.

The task instructs children that a bee is going to help them win points. At the beginning of each trial, children are asked to choose the number of flowers they want to put on the screen for their bee to stop at. For each flower the bee stops at they win one point. It is also explained to the children that the more flowers the bee stops at, the more tired it gets and the longer it takes to fly to the next flower. This set up is intuitive, meaningful and easy to grasp. Instead of being offered a binary choice (one point now vs two later), children are being prompted to engage with the task and set the magnitude of the rewards they desire themselves (children are offered 10 trials and are asked to choose between 1 and 7 flowers per trial).

The Bee Delay task also offers children the option of aborting trials early. Children have a stop button they can use during the waiting periods if they think one of their bees has become too slow. If they press it, they can move on to a different trial and start over while retaining the points they have won before aborting the trial. This feature of the task increases the accuracy with which delay time is measured by enabling children to adjust the overall delay time to the exact level they are comfortable with.

The task is also novel in that it gives three measures: a. an estimate of the child's preferences at the start of the trial in terms of rewards desired and achieved, b. the actual level of delay tolerated during the task, and c. an index of delay fatigue/frustration by measuring the discrepancy between the targets chosen at the beginning of the trial and the targets achieved by the end of it. A more detailed description of the Bee Delay task, its administration procedure and its outcome variables can be found in Appendix A.1.

4.3.1. Initial validation study of the Bee Delay Task

4.3.1.1. Participants

Sixteen typically developing children (9 boys and 7 girls) attending the reception year (Year R) at school were recruited through their school to help validate the Bee Delay task. Their school was adjacent to the University of Southampton campus, had participated in similar research projects in the past and was contacted directly by the researcher. Children's ages ranged from 4.88 to 5.32 years (M = 5.11, SD = .14). Prior approval for the study was obtained from the Ethics Committee for the Faculty of Social and Human Sciences of the University of Southampton.

4.3.1.2. Methods

Description of the task. Please see Appendix A.1.

Other measures. The Teddies Task (Sonuga-Barke, Dalen, et al., 2003) is a computerised choice delay task for children of a preschool age. Children are asked to make 20 choices between a large delayed reward –a teddy positioned in the background of the screen holding two balloons which takes a long time (17 seconds) to walk to the front of the screen to "release" the balloons and award two points to the participant– and a smaller immediate reward –a teddy holding one balloon but positioned closer to the foreground of the screen and thus taking less time (1 second) to release the balloon and award one point. Children made their choice between the two teddies by clicking on one of the two teddies presented on the computer screen. The side of the screen where the front/larger teddy was presented was counterbalanced. The computer recorded the number of LL (larger teddy, longer delay) and SS (smaller teddy, shorter delay) choices made by the children. The dependent variable used in this study was the percentage of LL responses. At the end of each trial, children were allowed to choose as many stickers as the number of points they won from a selection of stickers (with points ranging from a minimum of 20 to a maximum of 40). Young children's scores on the Teddies task have been shown previously to correlate significantly with performance on the CDT in a large sample of children aged between 3 and 5.5 years (Sonuga-Barke, Dalen, et al., 2003).

Strengths and Difficulties Questionnaire (SDQ, version T4-16 and follow-up). This is a behavioural screening questionnaire to be completed by teachers of children aged four to sixteen years. It includes 25 questions divided into 5 subscales (of five items each) measuring emotional difficulties, conduct problems, hyperactivity and inattention, peer problems and pro-social behaviours. The hyperactivity/inattention subscale includes items that tap inattention (2 items), hyperactivity (2 items) and impulsivity (1 item) as these are the three key symptom domains for ADHD. The psychometric properties and norms of this scale have been examined in large samples of children and have been found to be very satisfactory: internal consistency as measured by Cronbach's alpha: .73; test-retest reliability: 0.62 on average across subscales after four to six months (Goodman, 2001). For the complete SDQ scale, please see Appendices A.2 and A.3.

Quick Delay Questionnaire (QDQ, Parents and Teacher Form). This is a brief 10-item questionnaire that assesses delay related behaviours in children as rated by their parents or teachers on a scale ranging from 1 (not at all like him/her) to 5 (very much like him/her). Similar to the original self-report version of the scale (Clare, Helps, & Sonuga-Barke, 2010), it is comprised of two subscales, one focusing on delay and the other on delay discounting. For a more detailed description of the scale, also see section 4.4 of this chapter and Appendix A.4). The parent/teacher version of this scale was developed specifically for this thesis by adapting items included in the original QDQ and was validated using a pooled dataset from all the studies completed as part of this thesis (the full scale and scoring instructions can be found in Appendices A.5 and A.6).

4.3.1.3. Procedures

Children who agreed to participate in this study were asked to complete the Bee Delay task and the Teddies task in the "computer corner" of their classroom (Time 1; T1). The tasks were administered on a research laptop computer and not the school computers. Two weeks later (at Time 2; T2) the same children were asked to complete the same tasks again. A 10-trial version of the Teddies task was used instead of the original 20-trial version at T2. The children's teacher was asked to fill in the SDQ and the QDQ scales at T1 and only the QDQ scale at T2.

4.3.1.4. Analytic plan

An investigation of the Bee Delay scores distribution at T1 was conducted to rule out the existence of floor effects in this sample. T1-T2 correlations were calculated to establish the scores' stability over time. Finally, the correlations between Bee Delay scores, the main index of a neuropsychological choice delay task (Teddies Task) and the two behavioural rating scales were assessed in order to determine the task's convergent validity.

4.3.1.5. Descriptive statistics and score distributions

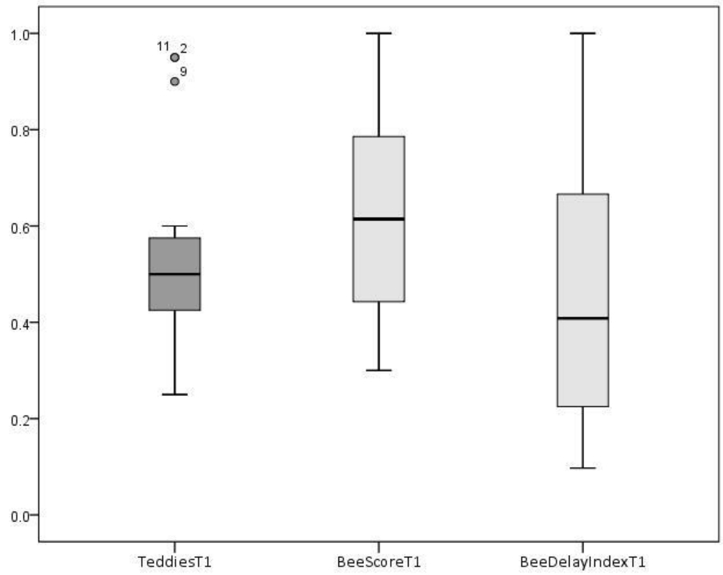
On average, children achieved 4.4 flowers (out of a possible 7) and waited for 41.47 seconds (out of a possible 90) per trial (Table 4.1). These indices were also calculated as ratios so they are directly comparable to the main outcome variable of the Teddies task, which calculates the percentage of LL responses. The frequency distributions of the two main indices of the Bee Task were found to be normally distributed using the Shapiro Wilk statistic (p = .268 for the Mean Flowers Score and p = .052 for the Mean Delay Index); the distribution of the Teddies scores was not (p = .004). Skewness and kurtosis values are also very low (<1). Box plots of these measures can be found in Figure 4.1.

Table 4.1. Means and standard deviations of all measures at T1 and T2
Test retest correlations (Pearson's, two-tailed) are also presented in this table.

	T1	T2	2 week test-retest Bivariate correlations
	M (SD)	M (SD)	
Teddies *(percentage of LL responses)*	.55 (.21)	.58 (.22)*	.92**
Bee Mean Flowers Score *(mean number of flowers achieved per trial)*	4.42 (1.55)	4.6 (1.44)	.80**
Bee Score *(flowers achieved as a percentage of maximum number of flowers offered per trial, i.e. 7)*	.63 (.22)	.66 (.21)	.80**
Bee Mean Delay *(mean time waited per trial)*	41.47 (26.51)	45.40 (26.64)	.83**
Bee Delay Index *(mean time waited as a percentage of mean time needed to achieve the maximum score, i.e. 90 sec)*	.46 (.29)	.50 (.30)	.83**
Discrepancy Index 1- Flowers *(Flowers achieved/ Flowers chosen)*	.97 (.10)	.95 (.07)	.03

	T1	T2	2 week test-retest Bivariate correlations
	M (SD)	M (SD)	
Discrepancy Index 2- Time *(Actual Time waited / Time needed to achieve all flowers chosen)*	.95 (.15)	.89 (.14)	.25
QDQ *(Total Score, range 0-40)*	18.38 (10.02)	12.75 (10.38)	.93**
SDQ *(Total Score, range 0-40)*	13.19 (7.24)	n/a	n/a
SDQ-Hyp *(Hyperactivity score, range 0-10)*	4.81 (3.31)	n/a	n/a

Figure 4.1. Boxplots of the two main indices of the Bee Task (Bee Score and Bee Delay Index), in comparison to the main index of the Teddies task.
All three indices measure performance as a percentage of a perfect score.

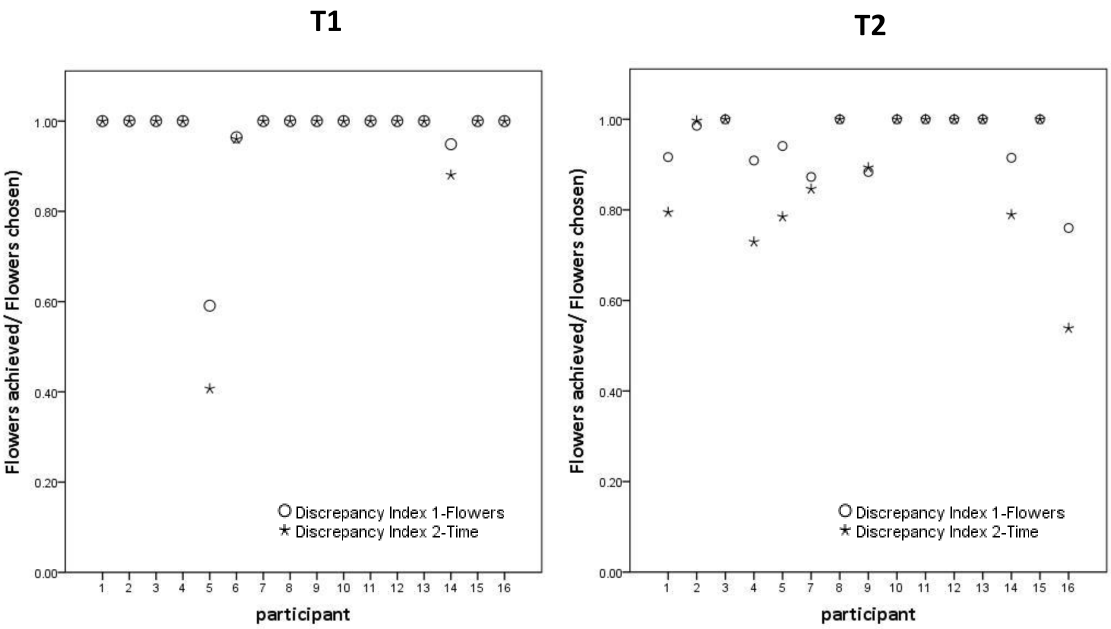


4.3.1.6. Test retest reliability

Test-retest reliability coefficients are also presented in Table 4.1. The stability of most indices was excellent, with the exception the two discrepancy indices. An analysis of performance by participant indicated that more children failed to achieve their targets at T2, reflecting possible increasing frustration with the task (Figure 4.2) and offering an explanation for the low test retest correlations. The 10-trial version of the Teddies was also found to correlate very well with the 20-trial version of the same task. This is a finding with practical implications, as it justifies the potential use of either of the two versions in future studies.

Figure 4.2. Discrepancy performance per participant at T1 and T2

A score of 1 indicates perfect agreement between the targets set at the beginning of each trial and the targets achieved. Discrepancy Index 1- Flowers: flowers achieved/ flowers chosen. Discrepancy Index 2- Time: actual Time waited/time needed to achieve all flowers chosen.



4.3.1.7. Convergent validity

Table 4.2 presents the correlations between a main index of the Bee Delay task, the percentage of LL responses children chose in the Teddies task and the ratings of teachers.

Table 4.2. Bee Delay scores and their correlations to a choice delay task and two teacher rating scales

		1	2	3	4	5
1	Bee Score at T1					
3	TeddiesT1	.53*				
4	QDQTotalT1	-.47	-.36			
5	SDQTot	-.49	-.39	.85**		
6	SDQHyp	-.34	-.34	.88**	.92**	

Note: Bee Score: flowers achieved as a percentage of maximum number of flowers offered per trial; QDQ: Quick Delay Questionnaire; SDQ: Strengths and Difficulties Questionnaire

The Bee Task positively correlated with the Teddies task and, as expected, negatively correlated with teacher QDQ and SDQ total and hyperactivity ratings. Higher scores on the Bee Task denote increased delay tolerance and would be expected to be associated with fewer delay related and behavioural problems. The latter negative correlations, however, albeit of moderate magnitude, failed to reach statistical significance. Also worthy of note are the high correlations of the QDQ with the SDQ, again in line with expectations.

4.3.1.8. Discussion

The aim of this study was to introduce a novel delay task designed specifically for children of a preschool age and to provide preliminary data on its psychometric properties. Its findings, as described above, provided initial evidence for satisfactory reliability, validity and age sensitivity of the task. The Bee Delay Task was shown to be a more sensitive and more flexible index of delay tolerance compared to existing choice delay tasks, as it provides a direct estimate of the total amount of time young children are willing to wait for delayed rewards without requiring an explicit understanding of the trade-off between reward magnitude and delay. The Bee Delay task rates delay choice and computes a continuous index of delay time. Given that the choice it offers children is not binary (children are asked to choose between 1 and 7 flowers per trial) it does not offer an estimate of LL preference directly, but enables children to pick the exact place they want to be at on the SS to LL continuum.

4.4. Adapting the Quick Delay Questionnaire for use by Parents and Teachers to rate preschool children

The original Quick Delay Questionnaire (QDQ) is a short, 10-item self-report questionnaire designed to assess delay related behaviour in adults (Clare, et al., 2010). It comprises of two subscales, tapping two distinct dimensions of delay intolerance: a. *delay aversion* focusing on positive or negative responses to waiting, which included items such as “I hate waiting for things” and “I feel relaxed waiting for things” and b. *delay discounting* rating the consideration of long term outcomes in relation to short-term outcomes, with items such as “I often give up on things I cannot have immediately” (see Appendix A.4). The scale has been shown to have good internal consistency, good test-retest reliability and to be moderately associated with ADHD (Clare, et al., 2010). The parent/teacher version of the scale was developed specifically for this thesis. It asks caregivers of preschool aged children to evaluate children's delay related behaviour along the same two dimensions (i.e. delay aversion and delay discounting) and on a scale ranging from 1 (not at all like him/her) to 5 (very much like him/her). Raters' responses are then shifted to a 0-4 scale before scoring, resulting in the QDQ total scores ranging from 0-40 (see Appendices A.5 and A.6).

4.4.1. Item content and development

All items were created primarily as adaptations of the original QDQ items. These had been selected from a larger pool of 43 items through a process of examining the item content and fit within the scales, as well as the factor structure of the scales (Clare, et al., 2010). One item of the current scale was revised to reflect behaviours of younger children, rather than of adolescents or adults: "Having to wait for things makes me feel stressed and tense" was revised into "Fidgety and restless when having to wait for things". The readability of the scale was assessed using an online readability calculator of the Flesch-Kincaid Reading Ease score (url: http://read-able.com/), which takes into account word and sentence length. This analysis revealed a Reading Ease score of 70.4 for the whole scale. This is interpreted to mean that the scale would be of easily understood by 13 to 14 year olds and thus of "average difficulty" and appropriate for use with most adult readers. This is according to a classification system introduced by the US Department of Health and Human Services (USDHHS), in the absence of relevant UK guidelines (Edmunds, Barry, & Denniston, 2013).

4.4.2. Validation study of the QDQ teacher rating scale

4.4.2.1. Participants. Data on 71 typically developing children were pooled together from three separate studies (the two feasibility studies presented in Chapter Five and the study piloting the Bee Task, presented above in section 4.3 of the current chapter) to form the combined normative sample. The ages of the children of this pooled sample ranged from 4.38 to 5.72 years (M = 5.08, SD = .33). Cases with missing data were excluded list-wise from all analyses reported in this chapter and for this reason the sample sizes may vary (all Ns are reported per analysis).

4.4.2.2. Analytic plan. The psychometric properties of the QDQ-Parent and Teacher version were examined by replicating and extending earlier analyses reported by the original developers of the scale. They included investigations of the internal consistency and reliability of the new version of the scale, the convergent validity of the scale with the Strengths and Difficulties Questionnaire (SDQ)-Teachers' version and the discriminant validity of the scale by comparing the QDQ scores of 24 children who met the clinical cut off for ADHD (from Study Three, see also section 6.4.3) to the QDQ scores of 24 children from the combined normative sample described above, matched on age and gender.

4.4.2.3. Internal consistency and test retest reliability. Cronbach's alphas were very high for the whole scale (10 items; α = .94) and for the Delay Aversion (DA) subscale (5 items; α = .93) and high for the Delay Discounting (DD) subscale (5 items; α = .86). To assess test retest reliability T1-T2 correlations were calculated for a subgroup of 29 children (n = 16 from the Bee Task pilot study and n = 13 from the waiting list group of the second feasibility study for whom T2 data has been obtained). The reliability estimates for both subscales (r = .88 for the DA subscale; r = .83 for the DD subscale) and for the whole scale (r = .89) were very satisfactory (all correlations significant at 0.01 level).

4.4.2.4. Convergent validity. The convergent validity of the QDQ scale was examined by comparing the correlations between the subscales of the QDQ and those of the SDQ (for more details on this measure, please also see also section 4.3.1.2). These are presented in Table 4.3. In line with expectations, QDQ subscales were mostly associated with the Hyperactivity and the Conduct Subscales of the SDQ.

Table 4.3. Correlations showing the relationships between QDQ and SDQ subscales

		1	2	3	4	5	6	7	8
1	QDQ Delay Aversion Scale								
2	QDQ Delay Discounting Scale	.85**							
3	QDQ Total Score	.97**	.95**						
4	SDQ Emotional Distress	.20	.38**	.29*					
5	SDQ Behavioural (Conduct) difficulties	.68**	.60**	.67**	.02				

		1	2	3	4	5	6	7	8
6	SDQ Hyperactivity and attentional difficulties	.80**	.81**	.83**	.25*	.72**			
7	SDQ Problems with peer relationships	.43**	.50**	.48**	.60**	.38**	.46**		
8	SDQ Total score	.72**	.78**	.78**	.68**	.69**	.85**	.79**	

Note: QDQ: Quick Delay Questionnaire; SDQ: Strengths and Difficulties Questionnaire

4.4.2.5. Discriminant validity. QDQ ratings of 24 children who met the SDQ Hyperactivity and Inattention clinical cut off in Study 3 were compared to ratings of 24 children from the combined normative sample matched for the child's age and gender. Details on the two groups' gender composition, mean age and QDQ scores are presented in Table 4.4. A MANOVA revealed significant differences between the group in terms of their QDQ scores (both subscales and total QDQ score), also evidenced by moderate effect sizes (partial eta squared also reported in Table 4.4). These findings demonstrate that children with ADHD were rated by their teachers as having significantly more delay related problems than children in the normative sample providing further support for the clinical sensitivity of the QDQ scale.

Table 4.4. Mean QDQ scores per group with means (SD) and statistical comparisons

	Clinical Cut Off Group (N = 24)	**Normative Group** (N = 24)	
	M (SD)	M (SD)	statistical comparison
Boys (% boys)	15 (62.5)	15 (62.5)	$\chi^2(1) = 1.0$, n.s.
Age (yrs)	4.75 (.46)	4.98 (.37)	$t(46) = -1.9$, n.s.
QDQ-Delay Aversion	14.25 (3.97)	7.25 (5.22)	$F(1, 46) = 27.35$, $p = .000$, $_p\eta^2 = .37$
QDQ-Delay Discounting	14.42 (3.59)	8.00 (4.43)	$F(1, 46) = 30.39$, $p = .000$, $_p\eta^2 = .40$
QDQ-Total	28.67 (7.27)	15.25 (9.48)	$F(1, 46) = 30.26$, $p = .000$, $_p\eta^2 = .40$

Note: QDQ: Quick Delay Questionnaire

4.4.2.6. Discussion. On the basis of the analyses presented above, the teachers' version of the QDQ scale was found to be reliable and to have internal consistency. It was also sensitive to clinical problems, specifically ADHD and CD. The QDQ scale was shown to associate strongly with SDQ subscales measuring hyperactivity, inattention and behavioural difficulties and to be able to differentiate a group of children meeting a clinical cut-off for ADHD from a group of typical controls. However, further research is needed to corroborate this finding as this analysis was done on a relatively small sample of very young children and used teacher ratings only.

Chapter 5: Development and feasibility testing of a novel delay training paradigm

5.1. Aims of the chapter

This chapter aims to describe the process of developing and further adapting the training protocol to be employed in this thesis. It also reports the results of two feasibility studies which examined the training efficiency of the task in two population samples of typically developing children.

5.2. Main features of the Waiting Game (WG) training protocol

According to the intervention model put forward by Sonuga-Barke (2004), interventions aiming to restructure delay preference should have two targets: to strengthen the association between a response and its associated delayed reward outcome (a link that is assumed to be weakened in ADHD) and to expose children to instances of easy-to-achieve, rewarded delay aiming to reduce the negative affect previously associated with delay.

The Waiting Game training programme was developed with these considerations in mind. More specifically the Waiting Game training task has the following features:

It uses short (i.e. manageable) delay periods spread over a large number of trials. The Waiting Game is based on a simple delay-to-reward task, in which children are exposed to a series of small rewards they can retrieve after brief delay periods. It was felt that having a moderately large number of trials, each offering small yet easy-to-achieve rewards was the best way to introduce children to the idea of practising waiting. This approach to training addresses both aims of Sonuga-Barke's (2004) intervention model: practicing on large numbers of trials should strengthen the basic association between the experience of waiting (however briefly) and of receiving a reward (however small) at the end of the waiting period. Brief delays also increase the likelihood that children's waiting will be successful (i.e. rewarded), which should reduce the negative affect they normally associate with waiting.

It is adaptive. The lengths of the delay periods are adapted to match individual children's initial abilities and learning rates during the training. Importantly, delays start off at manageable levels and they are continually adjusted upwards (or downwards, if necessary) as children practice on the task and their delay tolerance improves (or worsens). This adaptive feature ensures that children work at the limit of their abilities at all times, and that the waiting periods are neither too challenging nor too easy for them.

It is game-like. The task takes the form of a game where the particular rewards children have to wait for at every trial are determined by rolling a dice (for more details see Figure 5.3). The lengths of the delays are also varied (ranging from 5 to 30 seconds at the introductory level) and presented to the children in a pseudo-random order. This element of chance was introduced in order to make the task more unpredictable and thus less monotonous and more engaging. This is essential as children are expected to complete a fairly large number of trials while training. But please note all trials still had a certain outcome, with no risk involved. For the same reasons, the training schedule is not intensive: it includes five training sessions, administered over five days, with each session lasting no longer than 15 minutes in an effort to keep children interested and challenged throughout.

It is implicit. No instructions or feedback are given to the children in terms of what they can and cannot do during the delay periods. However, an observational coding system is implemented during the delay intervals in order to identify the behaviours children engage in while they are waiting and to capture changes in waiting behaviours over the duration of the training.

In the remainder of this chapter, the results of two feasibility studies are presented. The first of these studies used a one-group design to test the efficiency of the WG task. Its main objective was to demonstrate that young children's delay tolerance would indeed increase while they practiced on the WG task. As expected, results from this study lead to further adjustments of the training protocol. The second feasibility study used a three-group test-retest design to investigate the efficiency of the new, adjusted WG training protocol compared against two comparison conditions. Both feasibility studies also explored training effects, i.e. improvements in children's performance on non-trained tasks and delay related behaviours, as rated by their teachers.

5.3. Feasibility study 1

5.3.1. Aims and hypotheses

This first feasibility study had four main aims: firstly, to evaluate the potency of our chosen reward scheme by assessing whether the rewards on offer motivate the children to follow through with the training programme; secondly, to test the efficiency of the WG task by showing that the times children can tolerate delays for increase as a function of training; thirdly, to use observational data to demonstrate that increases in delay tolerance are accompanied by measurable changes in waiting behaviour and to assess the validity of a novel observational coding system; and finally, to explore the question of the generalizability of training effects by looking into whether improvements in delay tolerance during training may generalise to other delay related behaviours.

It was hypothesized that children's delay times would increase as they practiced on the Waiting Game task and that their performance on an untrained delay task would improve after training. Making predictions in terms of children's waiting behaviour and how that might change over time is more challenging given the paucity of relevant experimental evidence. It was predicted, however, that improved delay tolerance might manifest itself as a reduction of overactivity and inattention symptoms, in line with Sonuga-Barke's dual pathway model (2003). According to this hypothesis, hypersensitivity to delay may lead to increases in activity and inattention levels during delay periods that cannot be escaped. It is not unreasonable to assume, therefore, that as children become less adverse to the context of a delay-rich situation during the training, their activity and inattention levels will drop. Similarly, teacher ratings of children's relevant behaviours were also expected to reflect similar improvements, such lower levels of activity and higher levels of attention.

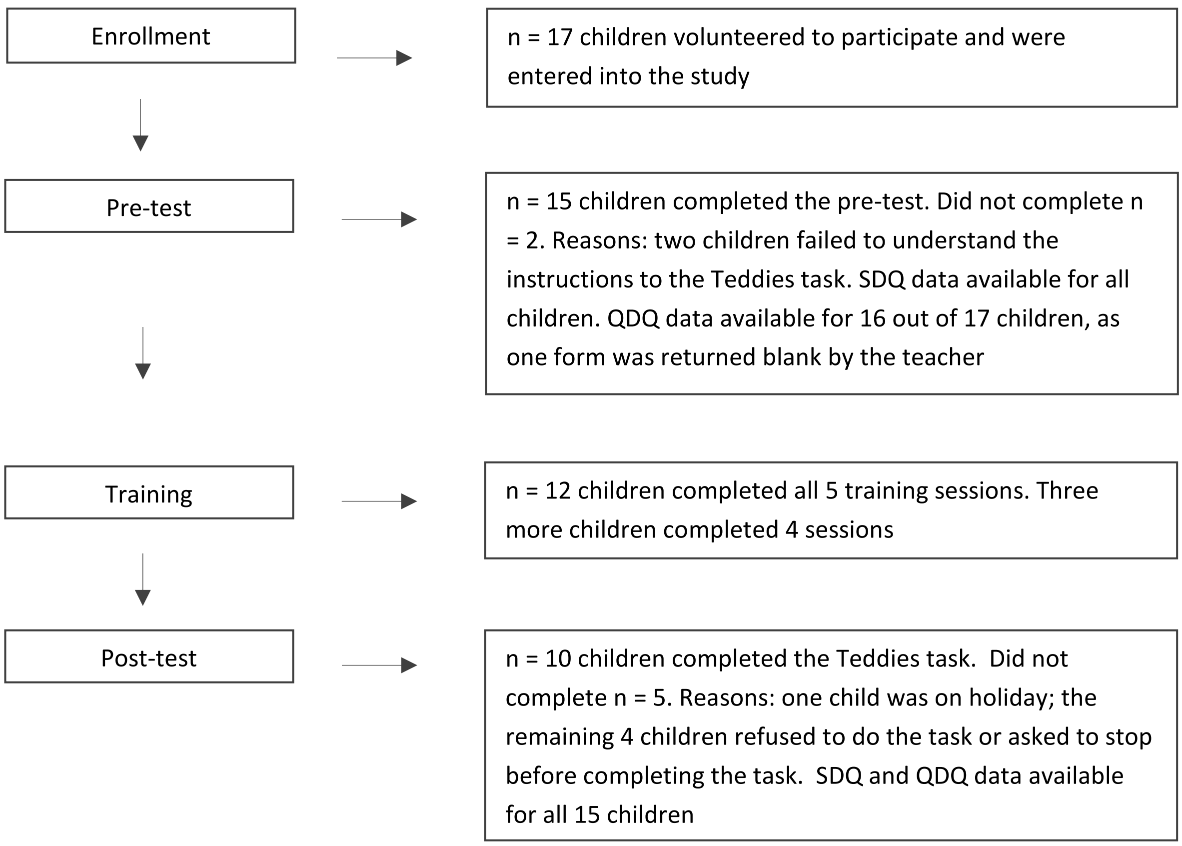
5.3.2. Methods

5.3.2.1. Participants

A total of 17 typically developing children (8 boys and 9 girls) attending Year R were entered into the study (Mean age = 5.41 years, SD = 0.26, age range: 4.82 - 5.72).

Children were recruited through their schools. All children attending Year R were eligible to participate in the study unless they had a statement of special educational needs, a diagnosis of a pervasive developmental disorder or had English as an additional language and had been assessed to have poor comprehension of oral English. Two of the children failed to understand the instructions to the Teddies task during the pre-test session and were excluded from the study as a result. For more details on children's participation as the study progressed see also Figure 5.1.

Figure 5.1. Diagram of study participation



5.3.2.2. Design and procedures

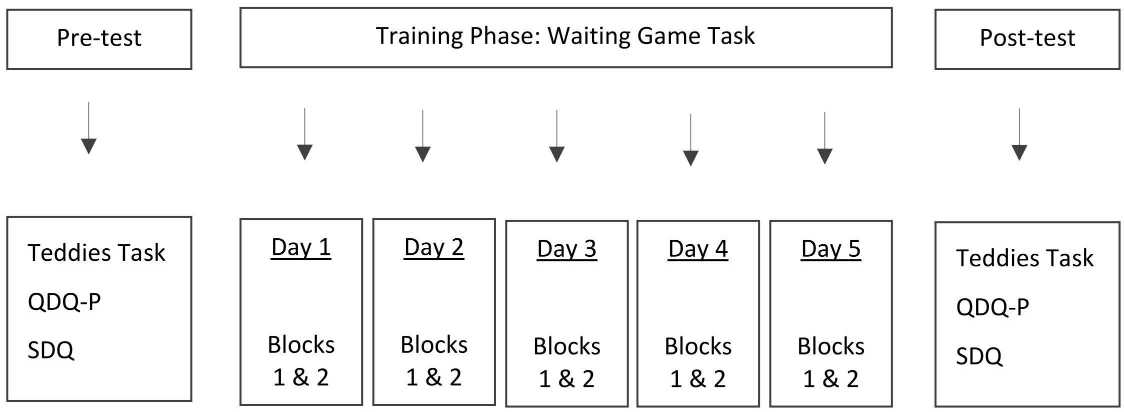
An initial invitation letter was sent out to the head teachers of primary schools in Southampton which participated in the School of Psychology Research Partnership Scheme. Two of the schools that expressed interest in the study were contacted again and a meeting was arranged between the head teacher of each school and the researcher. During those meetings the aims and the data collection methods of the study were discussed, as well as the inclusion and exclusion criteria for the study. Both head teachers gave their consent for the study to take place in their school premises and for eligible children to take part in the study on a voluntary basis. All Year R children were subsequently asked by their classroom teachers if they would like to volunteer to try out a new game. Teachers were instructed to explain to the children that they would be asked to play this game for a few times but they could at any point say no if they did not wish to do that anymore.

Children who volunteered to take part were asked to complete an initial assessment session (pre-test), where a computerised choice delay task was administered to them. Two behavioural rating scales were completed by teachers before the pre-test and after all children had participated in the post-test. Children who completed the pre-test were invited back for six additional sessions (five training sessions, and one post-test session, see also Figure 5.2). All sessions were administered on separate days over the course of two weeks and took place in a quiet room in the children's school. Prior approval for the study, which included a small pilot phase described in 5.3.2.3, was obtained from the University of Southampton, School of Psychology Ethics Committee.

5.3.2.3. The Waiting Game training task

Description of the task. As explained above, the Waiting Game training task was developed specifically for the present study. It is loosely based on the Cookie Delay Task (CDT)(Campbell, et al., 1982), which is a widely used delay of gratification type task for pre-schoolers (Sonuga-Barke, Dalen, et al., 2003). The original CDT involves placing an edible treat, such as a cookie, or a raisin, under a transparent cup and asking children to wait for a signal before they can retrieve it. It includes eight trials of varying delay intervals (ranging from 5 to 30 seconds) and administered in a fixed order. The CDT has satisfactory psychometric properties as it has been found to be reliable over time and to be able to discriminate between hard-to-manage preschool boys and their peers (Campbell, Pierce, March, Ewing, & Szumowski, 1994).

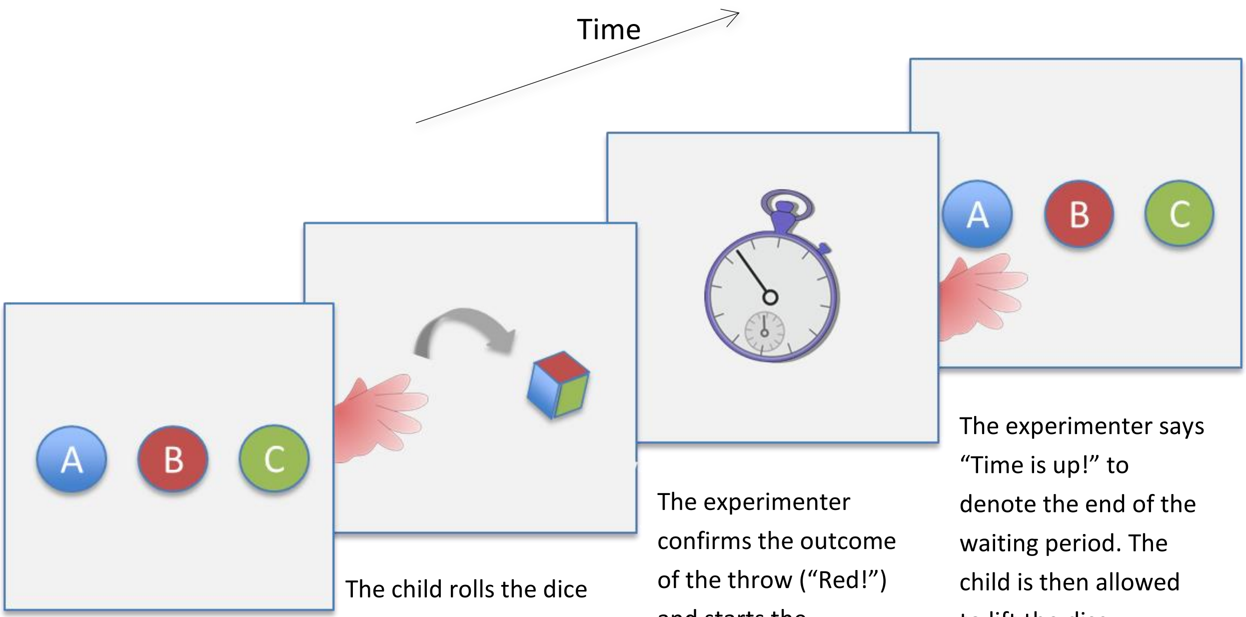
Figure 5.2. Diagram of the study's experimental design



In our version of the task, reward stickers were used instead of edible rewards in order to make the task more appropriate to use in educational settings. Stickers were of different forms, shapes and sizes, and as a result of varying degrees of desirability to the children (some were small and plain, some large and sparkly, for example). All children were given a blank sticker board at the beginning of their first session. It was explained to the children that they could collect all the stickers they won on these boards and that the boards would be theirs to keep after the study had ended.

At the beginning of each session, the experimenter set up three reward options by placing a random selection of stickers under three colour-coded transparent discs. Children were then given the following instructions: "*In this game you can win the stickers I have put under these circles. You can roll the dice to find out which circle to lift. The sticker under that circle is yours, you can have the sticker and you can stick it on your sticker board in any way you want…BUT in this game I want you to wait until I say "now" BEFORE you can get the sticker. Let's try."* The children were then encouraged to roll a three-colour dice to select one of the three options. A delay interval was imposed after the experimenter had confirmed the outcome of the dice roll (e.g. "Red!") and before the child could retrieve that reward. All three rewards were visible throughout the delay. When the end of the delay period was signalled by the experimenter, children retrieved the sticker they had won and proceeded to stick it on their sticker board straight away (please also see Figure 5.3 for more details on the sequence of events during a trial). Children who failed to wait until the end of a delay period were marked down (please find more details on the scoring procedure below) but they received no feedback on the outcome of the trial and were allowed to keep the stickers they retrieved early. One practice trial was administered and the instructions repeated at the beginning of each training session to make sure all children were familiar with the procedure. All children who completed the study were also awarded an achievement certificate on the last day of testing.

Figure 5.3. Sequence of events during a trial of the Waiting Game task.



Training protocol. Children completed five training sessions over a period of two weeks. Two blocks of eight trials were administered in each session, mirroring the structure of the original Cookie Delay task but also doubling its number of trials. This was done so that children’s motivational levels could be assessed by comparing their performance on the first and the second part of each session. The lengths of delay intervals used on the first day of training were the same as the ones used in the original CDT. On subsequent training days, children were moved on to more or less challenging versions of the Waiting Game task, which involved longer or shorter delay intervals respectively, depending on their performance on the previous day. Table 5.1 shows in detail the durations and the order of presentation of the delay intervals for each of the five training levels children were moved through during training. All children entered the study at Training Level 1. The two blocks administered on the same day were always of the same training level.

Table 5.1. Duration (in seconds) and order of presentation of delay intervals per training level.

	Training Level				
Trial	0	1	2	3	4
1	5	5	5	5	5
2	5	15	20	25	25
3	5	25	30	30	30
4	5	20	25	30	35
5	5	30	30	30	30
6	5	5	10	10	15
7	5	30	30	35	40
8	5	10	10	15	20
Total	*40*	*140*	*160*	*180*	*200*

Scoring. Each trial was scored as 0 = not inhibited (when the child actually retrieved the sticker before the end of the delay interval), 1 = partially inhibited (when reward-oriented movement was observed during the delay interval, e.g. the child touched, leaned forward or even lifted the cover of the sticker but did not attempt to actually take the sticker) and 2 = fully inhibited (when no reward-oriented movement or activity was observed). The possible range of scores for each block was therefore 0 to16 with a high score indicating increased inhibition. The inter-rater reliability of this scoring system was examined in a small pilot study of six children (average age: 5.7 yrs , SD = 18.7 months) recruited through an e-mail sent to all psychology staff of the University of Southampton. Children whose parents volunteered to take part in this pilot were videotaped in their homes while they completed the first day of training (two level 1 blocks of 8 trials each). Children's performance was then re-scored by an independent second rater who was given the full scoring instructions included in the Waiting Game manual and no further training. The agreement between the two raters was satisfactory (Cohen's Kappa: .66).

Advancement Criterion. To progress from one training level to the next across sessions, children had to achieve at least 50% of the highest possible score. Thus, children were advanced to the next training level if they had an average score equal to or higher than eight across the two training blocks they completed on a day. If a child failed to fulfil this criterion they remained at the same a level the next time they trained. If they failed to achieve 25% per cent of the highest possible score (or a score of 4), they were moved down a level.

The observational checklist. This observational system coded the different types of behaviours children engaged in while waiting. Behaviours falling into the behavioural categories described below were coded during the delay intervals only using a whole-interval method (Suen & Ary, 1989). According to this method, a score of 1 was recorded once for the whole delay interval when a target behaviour occurred at any point during that interval. Conversely, the non-occurrence of target behaviours within an interval would lead to a score of 0 for the interval. This method is informative only in terms of the frequency of occurrence of target behaviours but not the duration of the observed behaviours. The observational checklist was originally designed as a time-sampling based protocol to be implemented by a second researcher who would be able to observe the children independently from the researcher who led the children through their training. Unfortunately, this arrangement was not possible to put in place and the checklist had to be implemented by the same researcher who administered the WG. The checklist was then adapted using the whole-interval method as this method of sampling allowed for a more efficient use of the observer's time.

Because waiting is a composite behaviour, consisting of a constellation of different elements, a range of target behaviours were classified for observation. The process of selecting which behaviours to observe should be viewed as a continuing developmental process, particularly in the case of developing observational systems to be used in applied or clinical settings (Kanfer, 1985). With this consideration in mind, an initial selection of possible target behaviours was made on the basis of the following two criteria: behaviours identified in the literature as consistently able to distinguish children with ADHD from comparison groups (Platzman et al., 1992) and behaviours that were likely to arise specifically because of our particular experimental set up (e.g. orientation of the child towards or away from the rewards). Target behaviours identified in accordance to these criteria were subsequently organised into the six behavioural categories included the first iteration of the behavioural checklist:

Restless motor activity (RMA). The child displayed gross motor movements such as tapping a foot, rocking a chair back and forth, fidgeting, etc.

Aimless motor activity (AMA). The child displayed aimless, waiting specific movements while waiting (touching hair, playing with clothes or jewellery, etc.). Sometimes these movements would accompany gross body movements. If this happened, both codes (RMA and AMA) were recorded. Ordinary movements such as changing body or hand positions or pointing and gesturing while speaking were not recorded as AMA, unless they were very repetitive or done in an exaggerated manner.

Reward-oriented motor activity (ROMA). Child leaned towards and touched or lifted up the covers or in fact took the reward stickers when inappropriate. Sometimes it was very hard to differentiate RMA/AMA behaviours from ROMA because the child displayed, for example, aimless movement while at the same time leaned towards the sticker covers. In such instances, all relevant codes were used. Note that if the ROMA code was used, then the trial was always scored as a partially inhibited (or as not inhibited, if the child actually retrieved the sticker).

Looks up or away (LA). Child looked up and stared at the experimenter or looked elsewhere and did not engage with the activity at hand (was off-task). If a genuine distraction occured, e.g. someone walked into the room or music was suddenly heard, and the child looked up for this reason, a note was made as follows: LA + distraction. These instances were not counted as instances of LA.

Verbalization (V). Any verbalization by the child during a waiting interval were coded as V. Examples included the child making on- or off-task remarks, asking questions, requesting feedback or help, etc.

Out of seat (OS) or out of view. Any observed instance in which the child left his or her seat (or was out of view in case the session was videotaped). If the child suddenly stood up, for example, this behaviour was coded as OS.

In addition to coding behaviours falling into the above six categories, the observational checklist protocol also allowed the observer to record unexpected or additional waiting behaviours on an ad-hoc and informal basis to help with the further refinement or enhancement of these categories.

5.3.2.4. Pre- and post-training measures

The Teddies Task. As in section 4.3.1.2.

Teacher questionnaires. The classroom teachers of the children who took part in the study were asked to complete the following rating scales to assess children's behaviour before and after they had completed their training: Quick Delay Questionnaire-Parents and Teachers Form (QDQ). As in sections 4.3.1.2 and 4.4. Strengths and Difficulties Questionnaire (SDQ, version T4-16). As in section 4.3.1.2.

5.3.3. Results

5.3.3.1. Data screening

Prior to analysis, all performance variables and teachers' questionnaire data were examined using IBM SPPS Statistics 20 for accuracy of data entry, the presence of missing values and fit between their distributions and the assumptions of the statistical analyses used in this section, according to data screening procedures described by Tabachnick and Fidell (2007).

All children in the study were able to understand and follow through with the training task and there were no withdrawals from the study. However, due to poor school attendance during the summer months and occasional refusals to complete the second training block only three children (out of twelve who completed the five training days) had complete data for both training blocks on all training days. Very frequently, children opted to stop for the day after the end of the first block when, according to protocol, they were asked by the experimenter if they wished to continue or not. Given the scale of the loss of data, the decision was made to use all available data up to the point when children left the training sessions, regardless of whether a full training block was completed or not. Performance data were subsequently calculated on the basis of the number of trials actually completed. This allowed the calculation of a training efficiency index which focused on children's performance while they persevered with the training without taking into account the loss of data due to the early termination of the training sessions (a separate perseverance index, based on children's completion rates, was also calculated for comparison purposes). Data from children who completed 4 out of 5 days of training were also included in analysis.

A similar procedure was followed for the Teddies task. Even though all children completed the Teddies Task at T1, five out of fifteen children failed to complete all twenty trials of the task at T2. To determine the nature of the missing data, cases with missing and non-missing values on the Teddies task at T2 were assigned to two different groups and t-tests were run to test for mean differences between the groups in the remaining outcome variables (SDQ and QDQ scores as rated by the children's teacher at T2). These tests indicated there were no differences between the groups for either the SDQ, $t(10.8) = -.74$, $p = .47$, $d = -.41$ or the QDQ rating scale, $t(13) = -1.46$, $p = .17$, $d = -.08$, supporting the assumption that the post-test Teddies data were missing at random. Following this information, data from children who had completed at least half of the total number of trials of the Teddies task were included in the analysis and performance was calculated on the basis of the number of completed trials as above. This procedure resulted in the inclusion of data from two children who had completed

12 and 16 trials (out of 20). For the rest of the cases the T1 data point was carried forward, to provide a conservative estimate of change and preserve the original sample size in line with the intention to treat concept (Gupta, 2011). Because the proportion of data substitutions was so high, all analyses were repeated using complete cases only and results are reported for both variables: Teddies ITT (intention to treat) and Teddies Complete (complete cases only) where appropriate.

To identify outliers, the entire data set was scanned for z scores values in excess of 3.29 ($p < .001$, two-tailed). Two outliers were identified using this method. Both were in observation variables (adjusted frequency of ROMA on Days 2 and 4 of training). The impact of these outliers was reduced by replacing the outlying values with scores one unit larger than the next most extreme score in the distribution in line with procedures outlined by Tabachnick and Fidell (2007). Normality checks on T1/T2 variables using the Shapiro-Wilk statistic indicated that all but three variables were normally distributed. These variables (SDQ Conduct Problems at T1 and T2 and QDQ Delay Discounting scores at T2) were moderately skewed and logarithmic transformations were applied to improve them. The picture was very different when normality checks were run on the training efficiency and the observation frequency variables: out of 35 variables, only 12 were normally distributed. Closer inspection of frequency histograms and normal probability plots revealed that the majority of these distributions departed substantially from the assumptions of zero skewness and kurtosis. This was attributed to the nature of the data sampling methods used: firstly, the fact that the durations of the training delay intervals were fixed resulted in high prevalence of tied values in the training efficiency variables; secondly, the fact the non-occurrence of behaviours was scored as zero led to a very high prevalence of zero values in the observation frequency variables. These two factors, considered alongside the small sample size of the study, led to the decision to use non-parametric statistics in the analysis of the training efficiency and observational data. Non-parametric procedures make no distributional assumptions; instead, they make inferences by testing observed distributions against hypothesized distributions.

5.3.3.2. Analytical strategy

Descriptive statistics were used to assess children's perseverance with the training, as well as their performance over time during the training sessions (both within each training day and across the five days of the training). To further explore training efficiency, the mean delay time achieved by the group on the first day of training was compared to that achieved on the last day, using Wilcoxon signed-ranked tests. A series of bivariate correlations were used to assess the validity of the observational checklist against children's delay performance. Descriptive statistics and Friedman's ANOVAs were used to investigate changes in waiting behaviour as measured by the observational checklist over the duration of the training programme. An exploratory analysis of potential training effects was conducted by analysing teachers' ratings and children's scores on the Teddies task before and after training using paired samples t-tests. A correlational analysis was also carried out to examine the relationship between the degree of improvement on the three pre-/post-training measures and performance during training.

Effect sizes are reported using Cohen's *d*, corrected for the dependence between means across conditions (Morris & DeShon, 2002), in the case of the paired-samples t-tests and the effect size estimate *r* for the non-parametric procedures. Effect sizes are interpreted as .1, .3 and .5 reflecting small, medium and large effect sizes respectively (Cohen, 1992). Because of the exploratory nature of the study emphasis was placed on minimising the likelihood of Type 2 error (accepting a false null hypothesis). For this reason the significance level (α) was set at .05 throughout and was not adjusted for multiple comparisons for this study (Rutherford, 2011).

5.3.3.3. Training efficiency (did children's waiting times increase during the training?)

The mean duration of all delay intervals in seconds (mean Delay Time) for which a fully- or partially-inhibited score was achieved was calculated per child, per block. In cases of incomplete training blocks, mean DTs were averaged across the number of trials per block that had actually been completed. As explained in the previous section this procedure resulted in the calculation of an index of training efficiency on the basis of children's performance while they trained. The group's mean completion rate per training session (the number of trials completed over the total number of trials offered per session) was also calculated to provide more information in relation to the children's perseverance with the training (Figure 5.4). This figure suggests a differentiation in children's performance with training efficiency gradually increasing during the training while at the same time perseverance as measured by completion rates dropped.

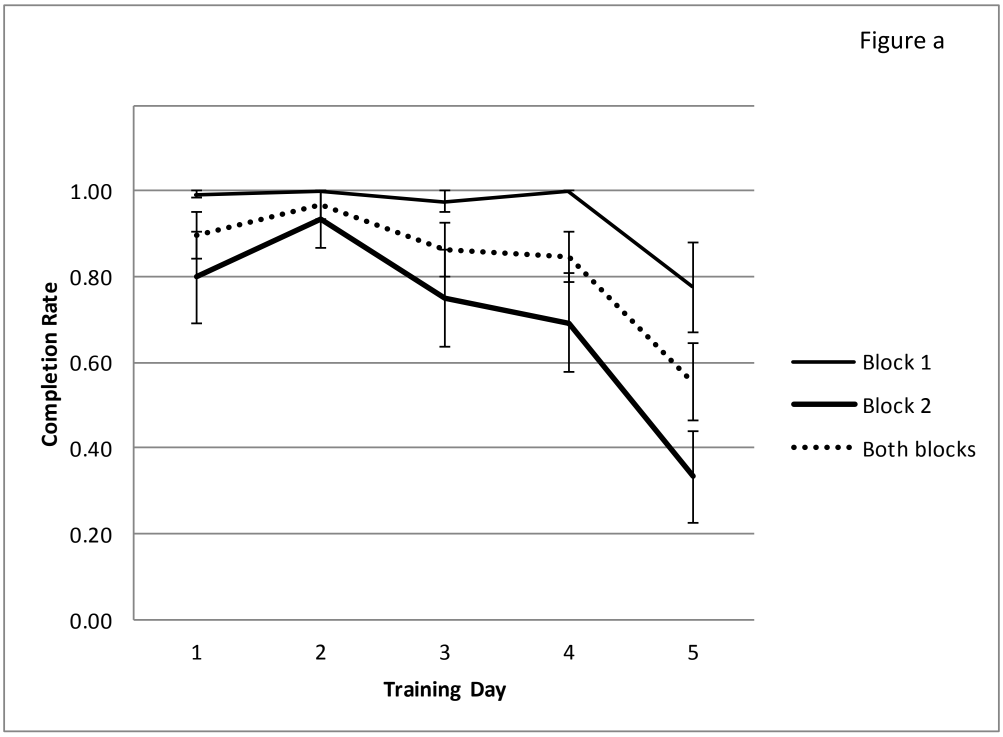
To further probe changes in performance over time, the mean DT achieved on Day 1 was compared with that achieved on Day 5 using Wilcoxon's signed ranked tests. Results indicated the group's mean DT did not improve significantly between the first (*Mdn* = 16.25) and the last (*Mdn* = 21.25) day of training, $z = -1.02$, $p = .31$ but a near-moderate effect size was detected $r = -.29$. The same was true when the mean DT observed in the first block of Day 1 (*Mdn* = 16.88) was compared to that observed in the first block of Day 5 (*Mdn* = 16.25), $z = -1.02$, $p = .31$, $r = -.29$. A smaller effect size was obtained when DTs between Days 1 (*Mdn* = 15.63) and 5 (*Mdn* = 22.50) were compared for the second training block only, $z = -.51$, $p = .61$, $r = -.19$.

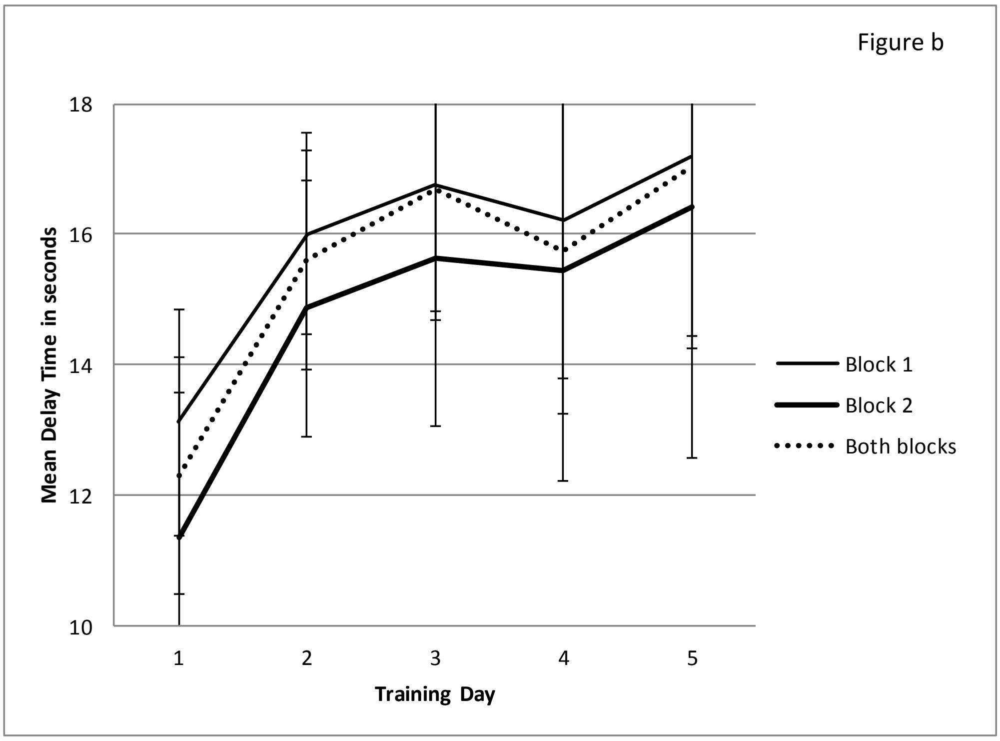
5.3.3.4. Changes in observed waiting behaviour (Did children's observed waiting behaviour change during the training period?)

Child waiting behaviour was coded within each of the six behavioural categories described in section 5.3.2.3. To correct for the fact delay intervals were becoming longer as children progressed through the training, and as a result more behaviours were likely to occur per delay interval, an adjusted frequency index was created for each behavioural category. This was calculated as the total number of times a particular behavioural category was observed during a training block divided by the total amount of delay time for which an inhibited or a partially-inhibited score was achieved. The adjusted frequencies per category are presented in Figure 5.5a. Among those, the categories with the most notable patterns of change across time have been isolated and are presented separately in Figure 5.5b. A seventh category is also included in these graphs for comparison purposes. This new category was created on the basis of the researcher's field notes which recorded instances of the children

Figure 5.4. Mean completion rate (Figure a) and mean Delay Time (Figure b) calculated per block across the five training days.

Error bars represent the standard error of the mean.

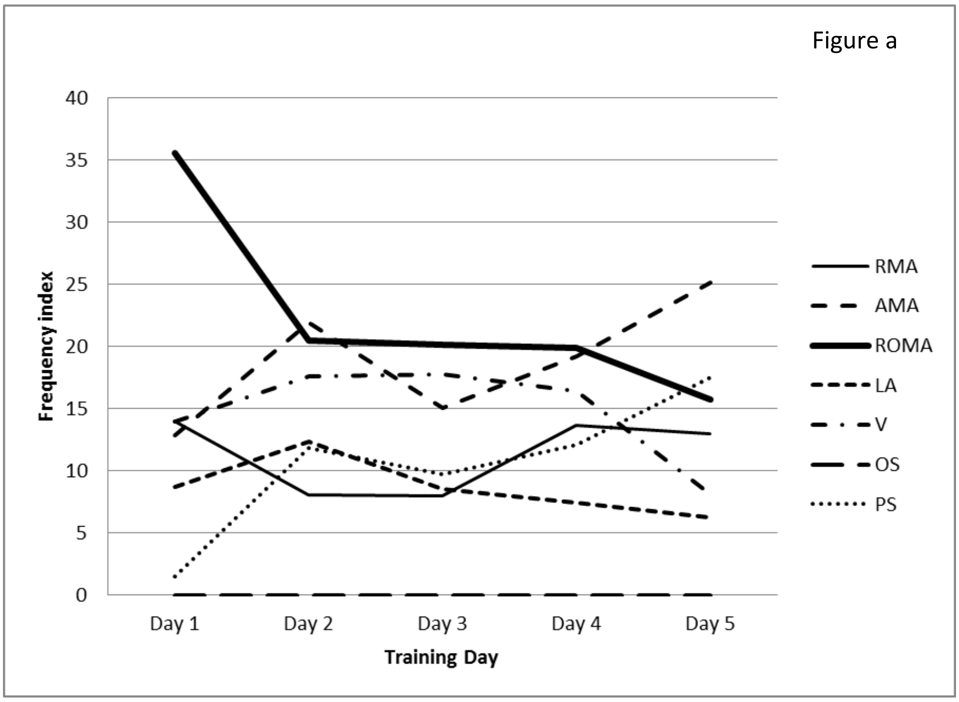


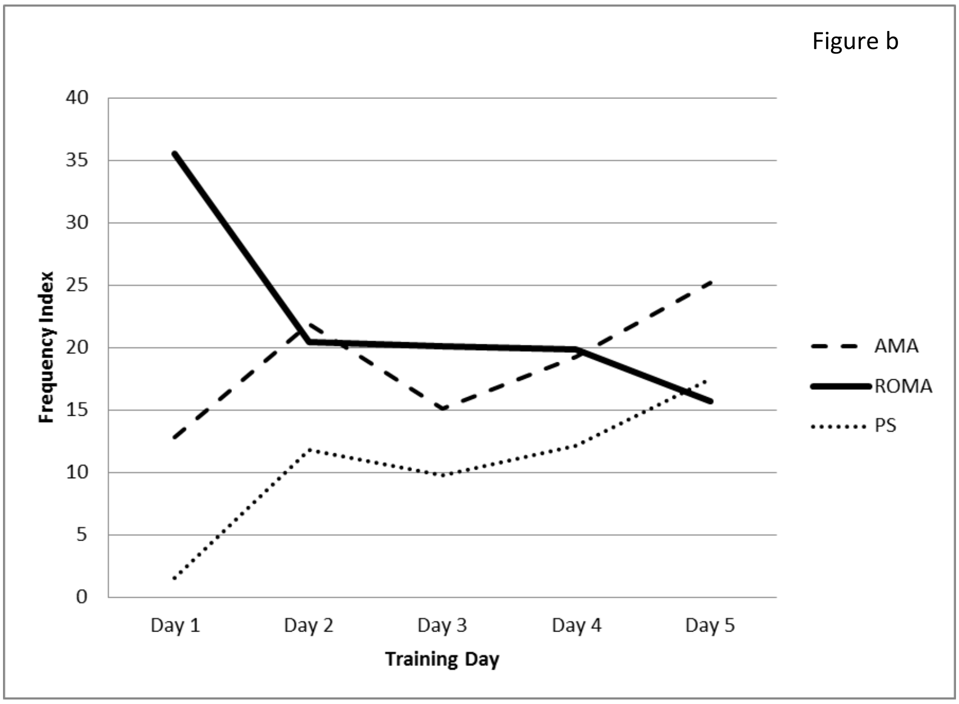


playing with the rewards they were winning during the training. This category was titled “Play with stickers” (PS). On the basis of these graphs, it appears that during the course of the training children are engaging in less and less reward-oriented activity and in more waiting-specific activity (aimless movements and playing with rewards already won).

Figure 5.5. Observed behaviours per category as a function of time (Figure a), with the most notable patterns of change across time presented separately in Figure b.

The behavioural checklist categories are: Restless motor activity (RMA); Aimless motor activity (AMA); Reward-oriented motor activity (ROMA); Looking away (LA); Verbalisation (V); out-of seat (OS) and the new category of Play with stickers (PS).The data presented in these graphs are pooled - with each training day acting as one data bin- and as a result estimates of variance are not computable in this case.





To explore these patterns of change over time a series of Friedman's ANOVAs were conducted on the adjusted frequencies of observed behaviours. Results indicated that none of the trends noted above achieved statistical significance but effects sizes were within the moderate to high range. More specifically, levels of children's restless motor activity did not change significantly over the course of the training, $\chi^2(4) = 2.47$, $p = .65$, $r = 0.31$ nor did levels of aimless motor activity, $\chi^2(4) = 5.30$, $p = .26$, $r = 0.57$; reward-oriented motor activity, $\chi^2(4) = 4.46$, $p = .35$, $r = 0.48$; looking away, $\chi^2(4) = 5.06$, $p = .28$, $r = 0.65$; verbalisation, $\chi^2(4) = 4.15$, $p = .39$, $r = 0.53$; or moving out of seat, $\chi^2(4) = .0$, $p = 1$, $r = 0$.

The validity of the observational checklist was estimated on the basis of how well its scores were related to the main training efficiency index of the study (criterion validity). More specifically, it was hypothesized that the children who achieved higher delay times on the first day of training would exhibit lower restless motor activity and be less likely to engage in inattentive behaviours (such as being off task or verbalising). The pattern of correlations between the mean DT achieved by the children on Day 1 and the adjusted frequencies per behavioural category are presented in Table 5.2. Contrary to expectations, results indicated that training efficiency was significantly correlated with waiting-specific activity (restless and aimless motor activity) and off-task behaviours (looking away) and negatively correlated with reward-oriented activity.

Table 5.2. Spearman's rho correlations between the mean Delay Time (mDT) achieved on Day 1 and the adjusted frequencies per behavioural category recorded on Day 1

The Out-of-Seat (OS) category was removed from the analysis because it was constant as no instances of OS were recorded for the duration of the training.

		1	2	3	4	5	6	7	8
1	Delay Time Day 1								
2	Completion Rate Day 1	.-16							
3	aRMA	.78**	.13						
4	aAMA	.64*	.14	.54*					
5	aROMA	-.91**	-.14	-.79**	-.39				
6	aLA	.63*	.31	.55*	.30	-.67**			
7	aV	.01	.34	.03	.26	-.06	-.05		

		1	2	3	4	5	6	7	8
8	aPS	.29	.20	.49	.08	-.38	.21	.10	

Note: Restless motor activity (RMA); Aimless motor activity (AMA); Reward-oriented motor activity (ROMA); Looking away (LA); Verbalisation (V); out-of seat (OS); Play with stickers (PS).

5.3.3.5. Training Effects (Did children's delay related behaviour improve after the training?)

Means and standard deviations at T1 (pre-test) and T2 (post-test) and results of paired samples t-tests for all pre-/post-training outcome measures are presented in Table 5.3. Even though improvements were noted for all measures between T1 and T2 at group level, none of the differences achieved statistical significance. Effect sizes were in the moderate range and ranged from 0.11-0.58.

Table 5.3. Mean Scores, Standard Deviations and paired t-tests results for all outcome measures at T1 and T2

	Pretest (T1)	Posttest (T2)						
	M (SD) [N]	M (SD) [N]	Raw Change	Percent Change	*t*	*df*	*p*	*d*
Teddies Complete	0.51 (0.09) [10]	0.57 (0.15) [10]	.06	.12	-1.19	9	.26	-.41
Teddies ITT	0.52 (0.08) [15]	0.55 (0.13) [15]	.04	.07	-.95	14	.36	-.22
SDQ Total Score	10.40 (4.88) [15]	9.67 (5.35) [15]	-.73	-.07	1.17	14	.26	.30
SDQ Emotional Difficulties	2.60 (2.67) [15]	2.67 (2.74) [15]	.07	.03	-.37	14	.72	-.11
SDQ Conduct Problems	0.34 (0.28) [15]	0.43 (0.41) [15]	.09	.25	-.43	14	.67	-.32
SDQ Hyperactivity	4.20 (2.08) [15]	3.67 (2.47) [15]	-.53	-.13	1.66	14	.12	.44
SDQ Peer Relationships	1.93 (1.98) [15]	1.60 (1.64) [15]	-.33	-.17	1.58	14	.14	.46
QDQ Total Score	18.64 (8.03) [14]	17.00 (8.28) [14]	-1.64	-.09	1.42	13	.18	.38
QDQ Delay Aversion	8.93 (5.20) [14]	8.00 (4.84) [14]	-.93	-.10	1.43	13	.18	.38
QDQ Delay Discounting	1.01 (0.15) [14]	0.76 (0.32) [14]	-.25	-.25	2.13	13	.05	.58

Note: Teddies Complete: complete cases included only; Teddies ITT: T1 values carried forward as per Intention-To-Treat protocol; SDQ: Strengths and Difficulties Questionnaire; QDQ: Quick Delay Questionnaire

Correlational analyses. A series of bivariate Pearson's correlations (α = .05, two-tailed) were used to explore the relationships between the pre-/post-training outcome measures at T1 (Table 5.4). In line with expectations, there was a strong correlation between the two rating scale measures at T1, $r = .66$, $p = .008$. The association between children's Teddies scores and teacher-rated SDQ scores at T1 was also considerable, but failed to reach significance ($r = -.45$ $p = .00$). In contrast, Teddies scores did not seem to correlate with teachers' ratings on the QDQ questionnaire ($r = -.12$ p = .66), which was contrary to expectations.

Table 5.4. Table of correlations between pre-/post-training measures at T1

		1	2	3	4	5	6
1	Teddies Score						
2	SDQ Total Score	-.45					
3	SDQ Hyperactivity Score	-.01	.50*				
4	QDQ Total Score	-.12	.66**	.52*			
5	QDQ Delay Aversion Scale	-.09	.63*	.51*	.97**		
6	QDQ Delay Discounting T1	-.14	.61*	.45	.92**	.78**	

Note: Teddies Score: percentage of LL responses; SDQ: Strengths and Difficulties Questionnaire; QDQ: Quick Delay Questionnaire.

5.3.4. Discussion

The results of this study showed that children's delay times improved during training, with moderate effect sizes obtained for comparisons specific to the first part of each training day. These gains were in stark juxtaposition to children's apparently falling levels of motivation to persevere until the end of each training day, which was evidenced by falling completion rates particularly during the second training block. Observational data suggested that over the course of the training children engaged in less reward oriented activity and in more waiting-specific and off-task behaviours. In addition, better delay times on the first day of training were significantly associated with this type of focus (away from the rewards at hand and towards off-task behaviours). Children's performance also improved on a delay aversion task and teachers reported reduced difficulties on the SDQ and QDQ scales post intervention. Again, these post-intervention gains did not reach statistical significance (but effect sizes varied from .11 to .58; Cohen's *d*). Despite the fact the experimental design of this study does not shed any light on whether these changes were related to the training (as they could have been entirely due to test-retest effects) it is still worthy of note that all recorded changes were in the expected direction.

The main objective of this study was to investigate the efficiency of the training task that was developed specifically for this study. Findings suggest the intervention was effective in a number of ways. To begin with, most of the children showed improvements in their ability to wait efficiently in the course of the training. In addition to that, the rewards offered during and at the end of training were effective in that there were no voluntary withdrawals from the study. Although the reward scheme used in this study successfully encouraged children to complete the study (since they could only take the rewards home with them on the last day of participation), it did not specifically encourage them to complete individual training sessions (as children could walk out of the sessions when they felt they had enough stickers for the day without any repercussions) or to strive for a good performance (as stickers were awarded at the end of each trial regardless of its outcome). It is possible that these manipulations contributed to one of the most remarkable findings of this study, namely the children's progressive loss of motivation to persevere until the end of the training sessions (only thirty three per cent of the children completed the second training block on the last day of training).

This pattern of results could be explained by reference to the resource depletion hypothesis (Muraven & Baumeister, 2000), which postulates that the ability to self-regulate relies on a limited resource that is depleted with use. According to this model therefore, children's effectiveness in inhibiting themselves should be expected to decline (and not to improve) shortly after practicing on a demanding inhibition task. Given the broad terms in which self-regulation has been defined by Muraven and Baumeister (2000), i.e. as any attempt to change one's behaviour in order to inhibit a dominant response, this model could be applicable in the case of delay training. What is more problematic is the two-task design paradigm that has most frequently been used in studies examining the model. Typically, participants are required to complete a first task that requires either a high or a low level of self-regulation followed by a second inhibition task. Poorer performance on the second task in the high self-regulation condition is interpreted as support for the model as it is hypothesized that engaging in the more demanding task depletes regulatory resources faster.

In the current study, there is obviously only one task. However, it is conceivable that the mid-session break creates the impression that a new "training episode" begins after the break. The first part of each training session could thus be assumed to serve as the demanding first episode –especially considering the adaptive nature of the training task which ensures children work at the limit of their abilities all the time– leading to a loss of perseverance and early exit during the second episode. Still, it would be difficult to determine whether it was simply the existence of a mid-session break which brought about the observed fatigue effects or other aspects of the training protocol as well, e.g. the overall length of each training session, the ratio of long-short delay intervals, the durations of delay intervals, the total number of training sessions, etc. Clearly, the resource-depletion hypothesis does not offer much guidance on these issues. Still, the model allows for long term improvements in self-control "strength" after practice, if the resource is allowed to replenish following a period of rest. This is a point not described in much detail by Muraven and Baumeister (2000) but it is an important one for the current study as it may help explain the fact children seemingly made gains in training efficiency from one block to the next despite overall increasing fatigue and loss in perseverance.

A second objective of this study was to use observational data to investigate whether increases in children's waiting efficiency were accompanied by changes in their waiting behaviour. Data obtained using the WG's observational checklist suggests that as children learned to tolerate longer waiting times, they also learned to switch their focus away from the rewards and towards other distractions. This finding was corroborated by a correlational analysis which revealed that better delay times on the first day of training were negatively correlated with reward oriented activity. In other words: high performers were able to disengage their attention from the rewards early in training. This pattern of results provides some independent evidence on the validity of the observational checklist. Despite the fact high performers did not engage in the types of behaviours we had predicted, it was still possible to correlate their performance during training with their behaviour while waiting in a meaningful way.

Children's performance on the untrained task is another issue to discuss. Improvements were noted, but as mentioned above, the tendency of some children to give up before reaching the end of the session continued during the post-test testing period and therefore affected performance on the Teddies task. This resulted in the loss of valuable data –particularly when participants dropped out before completing at least half of the trials. In addition to factors at play throughout the study, there are a few practical aspects of the way the post-test was set up that may have also affected children's performance on the Teddies task. Firstly, in line with the reward protocol followed during the training task, children were allowed access to the rewards they had won at the end of each trial. This was an integral part of the reward structure of the training task but it may not have worked so well in the case of the post-test. Stopping at the end of every trial in order to pick up the won stickers may have prolonged the perceived duration of all trials creating the impression they all lasted for about the same length of time and thus extinguishing the effect of choosing the shorter trials in order to escape delay. This may have led to children performing around the level of chance on the pre-test, which is more or less consistent with our data. Another practical consideration revolves around the issue of reward fatigue. It is important to note that children received the same rewards in the same experimental set up (same room, same researcher, etc.) during the pre- and post-tests and during training. It is reasonable to assume therefore that the novelty of the particular reward stickers must have worn off by the time of the post-test, leaving children with little motivation to see the last experimental session to completion.

5.3.4.1. Limitations

As a pilot, this study has some limitations. Firstly, the lack of a control group by definition restricts the type of questions this experimental design can address. Secondly, the small number of children who completed all phases of the study reduced the power necessary to explore these results in depth. Thirdly, this study used a normative sample to test an intervention designed to improve children with delay related difficulties; this fact alone may explain the existence of fatigue effects if one assumes that the training was too easy for children who do not find waiting situations particularly challenging.

5.3.4.2. Future Directions

Adaptation of the intervention. As highlighted above there are some aspects of the training task and training protocol that will need to be improved with a view to maximizing performance gains and limiting dissipation of children's motivation to complete the training to the extent possible. In future studies, special attention should be given to changes to the reward structure (for example, additional rewards should be offered when children manage to sustain delay until the end of the delay period), the structure of the training sessions (mid-session breaks should be removed and the overall length of the sessions reduced) and the pre- and the post-test set up (rewards offered then should be different to those used during training). To ensure that children are always faced with manageable delays, it is proposed that initial delay times be reduced slightly, as there were some children who found delaying challenging even on the first training day. For the same reason it is further proposed that the success criterion become stricter, to make sure that children do not move through delay levels too quickly.

Selection of control task/group. Because of the nature of this pilot study no control group was used. In the second feasibility study, where control conditions are used, careful consideration is given to the choice of an appropriate comparison training task. As discussed elsewhere (Markomichali, et al., 2009), often the best choice is a low-intensity version of the training task proper.

Selection of sample. Children participating in future studies should be selected on the basis of their difficulties in handling delay, to facilitate the ability to draw conclusions relating to a clinical sample. However, it was deemed premature to implement the WG training programme in a clinical sample without having first tested the adapted WG programme in a sample of typically developing children.

5.4. Feasibility study 2

5.4.1. Aims

This is a parallel arms randomised controlled trial designed to investigate the training efficiency of an adapted WG training protocol compared against two comparison conditions, a low intensity training condition and a waiting list control condition. A convenience sample of typically developing children was recruited for this study.

More specifically, and in accordance with the considerations discussed in the previous section, this second feasibility study had the three following aims: firstly, to evaluate the adapted WG task both in terms of training efficiency and in terms of children's perseverance to complete the training; secondly, to establish the inter-rater reliability of the adapted Observational Checklist and use observational data to confirm that increases in delay tolerance are accompanied by measurable changes in waiting behaviour and in particular by a shift of focus from the rewards at hand to off-task behaviours, in line with the findings of the previous study; and thirdly, to explore the issue of what constitutes an appropriate comparison training condition, by investigating differential training effects, i.e. performance gains in non-trained tasks and changes in teacher-rated and self-reported delay related difficulties, after training in a low-intensity version of the WG and after receiving no delay training at all.

We predicted that the adapted WG training protocol would lead to improved completion rates and gains in delay times across the training period. We also predicted that children would engage in less reward-oriented behaviours while waiting and instead turn their focus to other distractions, as suggested by the findings of the previous study. Finally, it was predicted that gains in delay related performance would be higher for the group who trained on the high-intensity version of the WG game, as opposed to the two comparison groups (the low-intensity group and the waiting-list group).

5.4.2. Methods

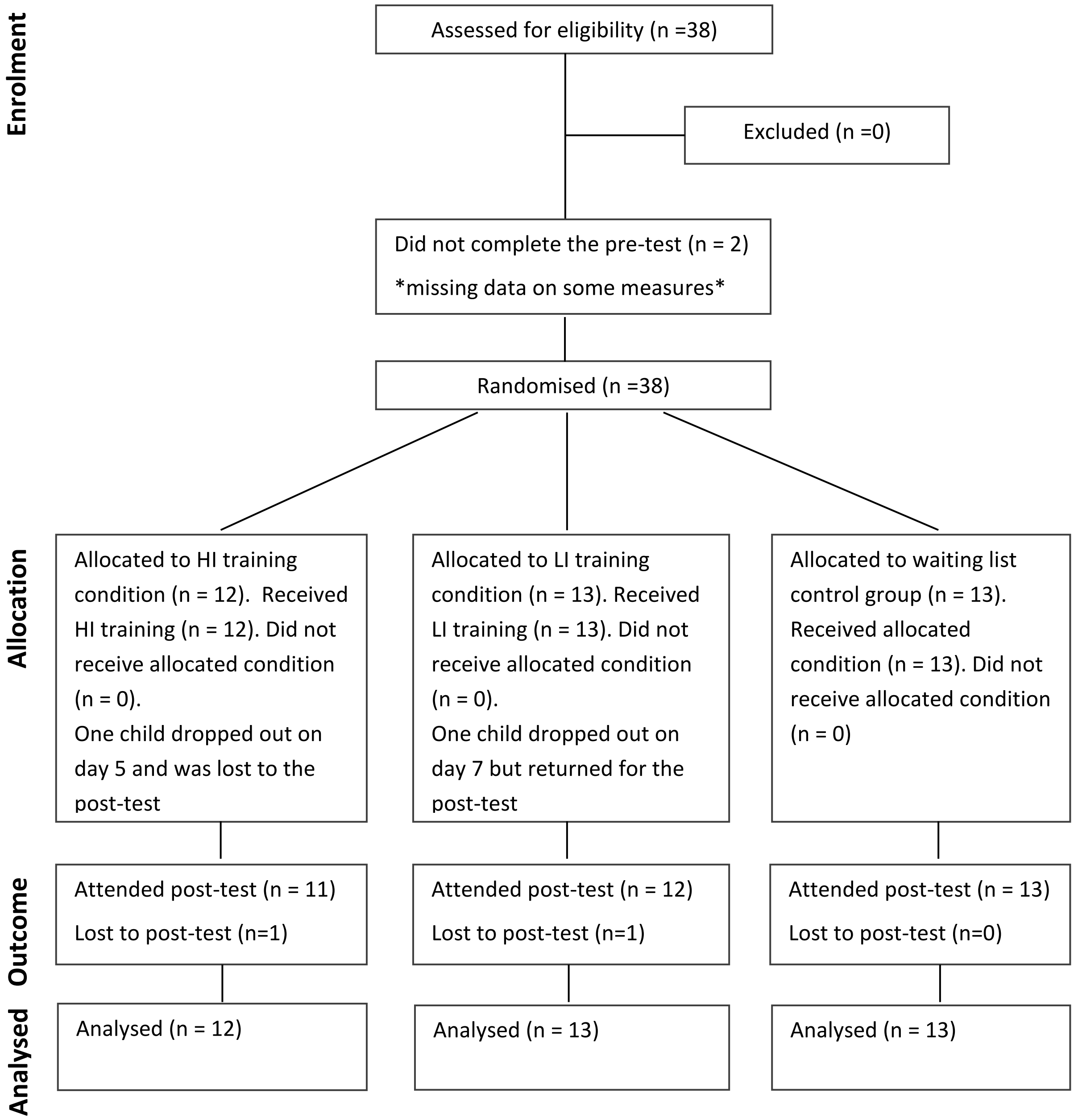
5.4.2.1. Participants, design and procedures

Participating children were recruited though their schools. As in the previous study, letters inviting local schools to take part in the study were sent to infant and primary schools in the Southampton city and surrounding areas. A research administrator then called the schools to inquire about possible interest in the study. Two schools expressed interest to take part and meetings were arranged with the head teacher of one school and the Year R lead teacher of the second school to discuss the practicalities involved in the study and procedures for obtaining parental consent. Both schools agreed to take part and to distribute information about the study to parents of Year R boys on behalf of the research team. Recruiting boys only was done in order to maximise the probability of achieving a sample enriched with cases of hyperactive and/or delay averse children. Informed consent was obtained from all the classroom teachers before they were asked to fill out questionnaires rating children's behaviour. Head teachers were offered the option of asking parents to opt out or opt into the study. For this study, one school asked parents to opt out if they did not want to participate in the study and the other school asked parents to opt into the study if they wanted their children to participate (12 and 26 children were recruited from each school, respectively). All procedures were approved prior to the start of the study by the Ethics Committee for the Faculty of Social and Human Sciences of the University of Southampton.

Exclusion criteria for the study were: a) a statement of special educational needs or previous diagnosis of a pervasive developmental disorder or b) in the case of children for whom English was not their first language, poor comprehension of spoken English as assessed by the teachers. Parental consent was obtained for 38 boys. These children were enrolled into the study and randomised into three comparison groups: high-intensity training, low-intensity training and waitlist control. Of the 38 enrolled children, 12 were assigned to the High Intensity (HI) condition, 13 to the Low Intensity (LI) condition and 13 to the waiting list (WL) control group (Diagram 5.6). Their ages ranged from 4.4 to 5.5 years (Mean age = 4.9, SD =.31). Randomisation was done in blocks of three to avoid disparities in the sizes of the three groups. The generation of the allocation sequences was done using the website www.randomiser.org which uses a JavaScript random number generator to produce customized sets of random numbers.

Children completed a neuropsychological assessment session before and after the training phase of the study (pre- and post-test). Three behavioural rating scales were also completed by teachers as part of the pre- and post-training assessment. All sessions, including assessment and training sessions, were held on separate school days over the course of three weeks and took place in a quiet room in the children's school. Assessment sessions lasted for approximately 20-25 minutes and training sessions about 10 minutes each. Children were offered breaks during the assessment sessions, but because the training sessions were so brief children were not offered breaks during the training sessions.

Figure 5.6. Consort Diagram



5.4.2.2. The adapted Waiting Game (WG) training task

As discussed in section 5.3.4.2 the results of the first feasibility study led to a number of adaptations to the WG training protocol and reward scheme. These are summarised in Table 5.6.

Table 5.5. Summary of modifications to the WG training task and protocol

Problem	Possible reasons why	Resulting adaptations
Low completion rate during training	No reward contingency in place for non-inhibited trials	Introduction of relevant reward contingency
	Delays becoming too long too quickly?	More training sessions, and stricter advancement criterion to ensure more gradual increases of delay times
Completion rate higher and performance gains lower during the second training block	Fatigue effects (resource depletion hypothesis)	Training administered in one block only
	Training sessions too long (too many trials)	Fewer trials
	Delay intervals too long	Shorter delay intervals
Children often off-task while waiting	Rewards offered opportunities for distraction	Rewards awarded at the end of session as opposed to the end of every trial
Low completion rate at the post-test	Reward fatigue	Different reward system in the pre/post assessments
Teddies performance around 50%	Awarding rewards after every trial acted as a post-reward delay?	Different reward system in the pre/post assessments

More specifically, the following changes were implemented:

Enhanced reward contingency: In this version of the task, an enhanced reward contingency was introduced. Children were asked to wait for a reward, but they were also instructed that if they waited successfully until the end of the waiting period, they could access not only the reward they waited for, but a second more valued reward as well. In a sense therefore the WG task was turned into a choice delay task in that children had a choice between exiting a trial early while still obtaining a small reward or waiting until the end of the trial to secure an additional larger reward. In order to access the larger rewards children were asked to choose an additional sticker at the end of every successful trail. These stickers were retrieved from a special box which contained very attractive, i.e. colourful, sparkly and uniquely shaped, stickers. The smaller rewards used in the normal course of the game were plain small smiley face stickers.

However, it needs to be emphasized that children were never explicitly given the choice not to wait. The exact instructions given to the children were as follows: "*In this game you can win the stickers I have put under these circles. You can roll the dice to find out which circle to lift. The sticker under that circle is yours, you can have the sticker and you can stick it on your sticker board in any way you want…BUT in this game I want you to wait until I say "now" BEFORE you can get the sticker. If you wait, then you can take a second sticker from the box with the special stickers. YOU CAN THEN KEEP BOTH STICKERS. Let's try."*

At the end of each trial, and depending on the outcome of the trial, the experimenter said: *"Good job/Well done! You can take this sticker now and you can pick a second sticker from this box because you waited until I said "Time's up!"* or *"Oh dear! This sticker is yours, but you cannot pick a second sticker this time because you didn't wait for me to say you could have the sticker."*

In line with the adaptations laid out in Table 5.6 children were asked to place all the stickers they won during the course of the game in a container they were given at the beginning of each training day, and were only allowed to use the stickers (i.e. stick them on their sticker boards) at the end of each training session.

Training protocol. Table 5.7 presents the new training schedule. Overall, the changes to the training schedule can be summarised as follows: The adapted WG task is now delivered in one block of 12 trials per day instead of two blocks of 8 trials each. However, the overall duration of the new training protocol is longer as it is implemented over eight days as opposed to five. This ensures a more gradual increase of delay times as the training progresses. Children are asked to complete a slightly larger number of trials overall, but the total delay time they are asked to endure per session is less (300 seconds on the last training day compared to 400 seconds on the last day of training of the previous feasibility study). All children start the training at level 1.

Table 5.6. Duration and order of presentation of delay intervals per training level
Delay times implemented in the original WG training task are included for comparison purposes.

	Training Level								
Trial	0	1	2	3	4	5	6	7	8
1	5	5	5	5	10	10	10	15	15
2	5	10	10	10	15	20	25	25	25
3	5	10	10	20	25	25	30	30	40
4	5	5	15	15	10	15	15	20	20
5	5	10	5	5	5	30	10	10	10
6	5	5	15	10	25	10	25	30	30
7	5	5	10	15	10	15	10	15	15
8	5	10	10	20	20	15	35	35	45
9	5	5	15	5	15	25	40	40	20
10	5	10	5	20	30	15	10	15	40
11	5	5	10	10	10	20	20	20	20
12	5	10	10	15	5	10	10	15	20
Total	*60*	*90*	*120*	*150*	*180*	*210*	*240*	*270*	*300*
Total Study 1	*80*	*280*	*320*	*360*	*400*				

Advancement criterion: The advancement criterion was stricter than in the first pilot study: 66% (or, in other words, a score equal to or larger than 16 out of a total 24). If this criterion was not met, children remained at the current training level. If they scored 8 or below (equal or lower than 33% of the highest possible score) they were moved down a level). Making the criteria stricter ensured children are not moved through the levels too soon.

Reward scheme: To avoid reward fatigue a completely new reward scheme was introduced during the pre- and the post-test assessment sessions. During these sessions children were awarded tokens for their performance, which they could exchange with small gifts at the end of the sessions. Depending on their performance children could exchange their tokens for either one or two gifts (if they had achieved more than 50% of the highest possible score, they could get two gifts. The cut-off point for being awarded two gifts instead of one was communicated to them at the beginning of the session. The gifts were stationary items pre-approved by the teachers (novelty pencils, erasers or notebooks).

Low intensity training control condition. Children allocated to the low intensity training condition trained on the same WG task and were offered exactly the same rewards as children training in the high intensity condition. The only difference between the two was the length of the delay intervals: in the low intensity condition children were asked to wait for five seconds only before retrieving their rewards with this interval being constant through the 8 days of training. Given the brevity of the training trials, no observational or training efficiency data were recorded for this condition (with the exception of completion rate data).

5.4.2.3. The observational checklist

The observational checklist was adapted in response to the findings of the previous study. More specifically, a new behavioural category was introduced, that of *distracted attention* (D), to capture instances where the child attempted to distract themselves during waiting. To adequately differentiate instances of Distracted attention from other behavioural manifestations, e.g. looking away, the remaining behavioural categories were further refined. The updated descriptions of all the behavioural categories can be found in Appendix A.7. In addition, all behavioural categories describing behaviours which could be regarded as off-task or as helping children distract themselves while waiting were collapsed into one aggregate category named Overall Distraction Index. This was calculated as the total count of behaviours falling into the following four categories: Aimless motor activity (AMA); Looking away (LA); Verbalisation (V) and Out of seat (OS). To establish inter-rater reliability a second independent scorer observed all children allocated to the HI training condition on their sixth day of training.

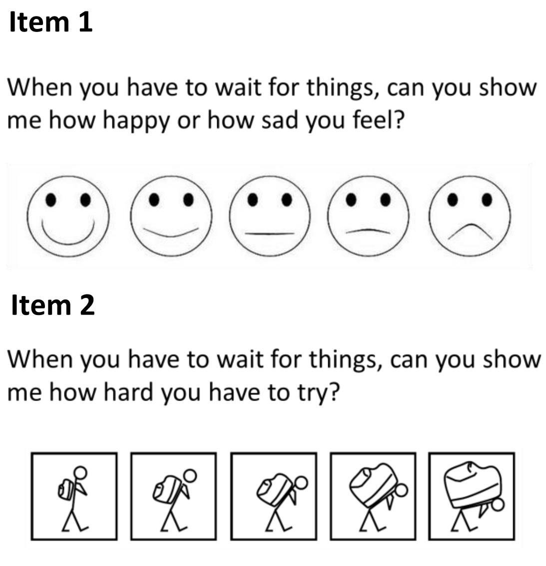
5.4.2.4. Pre-post-training measures

The Teddies task. As in section 4.3.1.2. This was selected as the primary outcome measure of this study in order to maintain continuity between the two feasibility studies.

The original Cookie Delay task. As in section 5.3.2.3. This was a task very similar to the one children trained on during the training phase of the study. It was selected as a measure of proximal improvements.

The Delay Aversion Self-report Scale for pre-schoolers (DASS-P). This is a self-report measure of delay intolerance developed specifically for this study. It comprises of two items: the first item uses a faces scale consisting of five faces with facial expressions ranging from happy to sad (Figure 5.7). Children were asked to point to the face which best describes how they feel when they have to wait for things. Their responses were scored from 0 (happy) to 4 (sad); the second item of the scale uses a series of pictures depicting a human figure carrying a rucksack varying in size from small and easy to carry to very large and difficult to carry. Children were asked to describe how hard they find waiting by pointing to the size of the rucksack which best describes how easy or how hard they find waiting. The scores on the second item also ranged from 0 (easy) to 4 (sad), making the total scale range from 0 (low delay aversion) to 8 (high delay aversion).

Figure 5.7. The two picture scales included in the Delay Aversion Self-report Scale for Pre-schoolers (DASS-P)



Teacher questionnaires. The classroom teachers of the children who took part in the study were asked to complete the following standardised rating scales to assess children's behaviour before and after the training phase of the study: Quick Delay Questionnaire-Parents and Teachers Form (QDQ). As in section 4.3.1.2. Strengths and Difficulties Questionnaire (SDQ, version T4-16). As in section 4.3.1.2. The Behavior Rating Inventory for Executive Function-Preschool Version (BRIEF-P). The BRIEF-P is a rating scale designed to measure executive functions in children aged two to five years. It consists of a single form that can be completed by different raters such as parents, caregivers or teachers. It measures multiple aspects of executive functioning including inhibition, set shifting, emotional control, working memory and planning. The BRIEF-P yields an overall index, the Global Executive index and three further composite indexes: the Inhibitory Self-Control Index (ISCI), Flexibility Index (FI), and Emergent Metacognition Index (EMI). The ISCI index, which is composed of the Inhibit and Emotional Control scales, is of particular interest to this study and will be included in the analyses. The psychometric properties of the BRIEF-P are very satisfactory: individual clinical scales have good internal consistency and convergent validity with a widely used measure of child psychopathology, the Child Behaviour Checklist (CBCL; Duku & Vaillancourt, 2013; Sherman & Brooks, 2010). The scale is also sensitive to symptoms of ADHD, with children diagnosed with ADHD being rated significantly higher than controls on all primary BRIEF-P measures (Mahone & Hoffman, 2007).

5.4.3. Results

5.4.3.1. Data screening

Prior to analysis all performance variables and teachers' questionnaire data were examined using IBM SPPS Statistics 20 for accuracy of data entry, missing values and fit between their distributions and the assumptions of the statistical analyses used in this section, according to data screening procedures described by Tabachnick and Fidell (2007).

In terms of missing data, the drop out and completion rates were much improved compared to the first pilot study. Two children failed to complete all the tasks at pre-test, but did not withdraw from the study and went on to complete most of the training sessions. One child dropped out from the high intensity group on Day 5 of training and was lost to post-test. With regard to this child's training performance, all available training data was used up to the point of completion in line with the procedure followed in the previous study and performance data were calculated on the basis of the number of trials completed on each training day. For all other variables, a missing data analysis was conducted. The assumption that data were missing completely at random was tested using Little's MCAR test. This analysis resulted in a chi square of 15.57 (*df* = 30; *p* = .99), which indicated that the data was missing at random (i.e., no identifiable pattern existed to the missing data). Following this result, the Expectation Maximization (EM) algorithm was used to compute maximum likelihood estimates of the missing values (Jelicic, Phelps, & Lerner, 2009). Another child dropped out from the low intensity training, but returned for the post-test. Since no training data were recorded from the low intensity group, this child's dropout did not result in any missing data.

Screening procedures for outliers and checks for the normality of variables were carried out separately for the three groups according to the protocol for screening grouped data outlined by Tabachnick and Fidell (2007). The entire dataset was scanned for cases with standardised scores in excess of 3.29 ($p < .001$, two-tailed) and no outliers were detected using this method. However, a large number of variables were found to be non-normally distributed when tested statistically to evaluate the significance of their skewness and kurtosis (44 out of 151 normality checks using the Shapiro-Wilk statistic were statistically significant, indicating the data were normally distributed within each group for about one third of all variables). This was attributed again to the nature of the sampling methods used (please also see section 5.3.3.1 for a relevant discussion) and the small sample size of the study, as many of the pre-/post-training outcome variables were found to deviate significantly from normality when assessed within groups. Non-parametric procedures were used to analyse the training and observational data as before, but given the robustness of the analysis of variance procedure (Wilcox, 2001) the decision was made to proceed with parametric analyses for estimating test-retest differences between the three groups, using repeated measures analyses of variance (ANOVA).

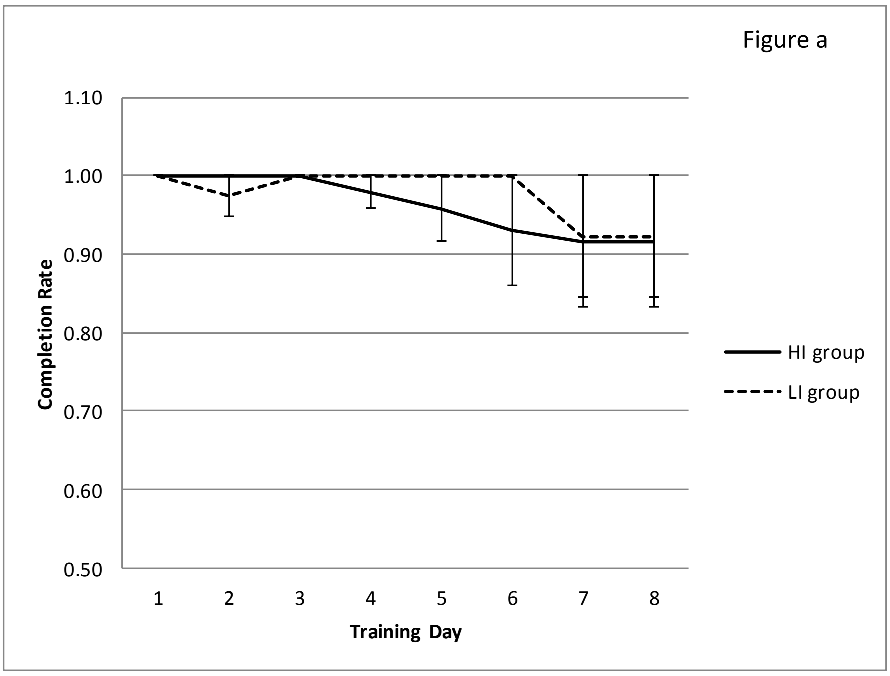
Analytical strategy. As in the previous study, descriptive statistics were used to assess children's perseverance with the training. To further assess the training efficiency of the group who underwent the high intensity training, the mean delay time achieved on the first day of training was compared to that achieved on the last day, using Wilcoxon matched pairs signed-ranked tests. Descriptive statistics and Friedman's ANOVAs were used to investigate changes in waiting behaviour over the course of the training as before. To assess consistency in coding between two observers, Cohen's kappa was calculated separately for each behavioural category. A correlational analysis (Pearson's, two-tailed) was used to explore the relationships between observed behaviours and children's training efficiency on the first day of training. Cronbach's alpha and a series of Pearson's correlations were conducted to explore the internal consistency and convergent validity of the newly introduced delay aversion self-report scale. A series of 3 (group) x 2 (time) repeated measures ANOVAs were conducted on all pre-/post-training outcome measures, with age added as a covariate to the repeated measures design, to investigate within- and between-group effects on children's test-retest performance. Effect sizes are reported using partial eta squared for the repeated measures ANOVAs and the effect size estimate r for the non-parametric procedures. Effect sizes are interpreted as .1, .3 and .5 reflecting small, medium and large effect sizes respectively (Cohen, 1992). As in the previous study, the significance level (α) was set at .05 throughout and was not adjusted for multiple testing because of the exploratory nature of the study and its emphasis on minimising the likelihood of Type 2 error (Rutherford, 2011).

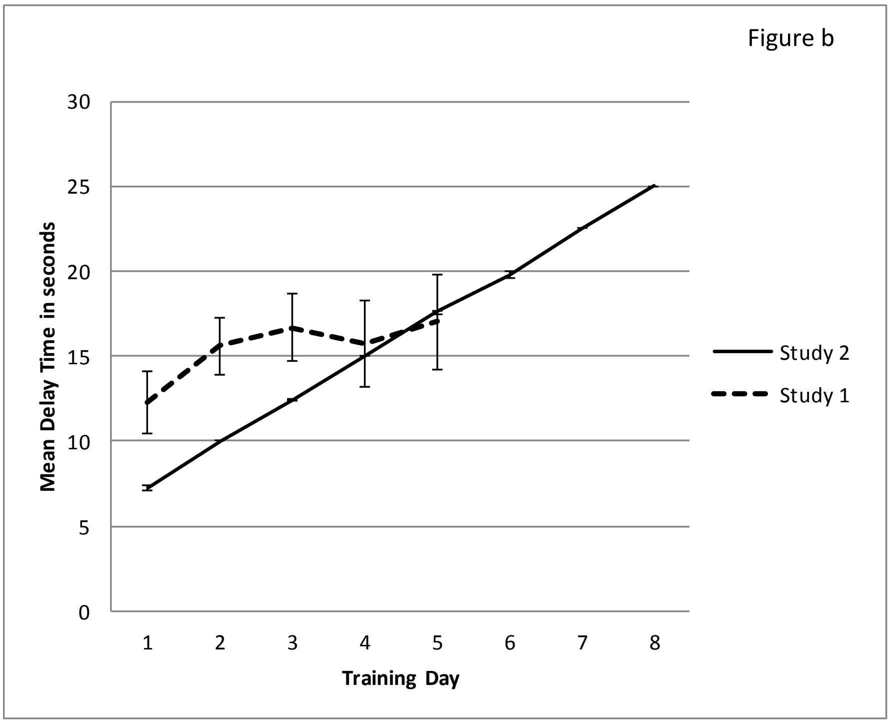
5.4.3.2. Training efficiency

As in Study One, the mean duration of all delay intervals in seconds (mean Delay Time) for which a fully- or partially-inhibited score was achieved was calculated per child per training day. The two training groups' mean completion rates per training session (the number of trials completed over the total number of trials offered per session) are also presented in Figure 5.8. This figure suggests that children's perseverance improved considerably under the adapted WG training protocol with both groups achieving a completion rate of more than 90 percent. Children's mean Delay Time reached 25 seconds per trial on the last day of training, which was the highest delay time they could have achieved according to the training protocol. When the mean DT achieved by the HI group on Day 1 was compared with that achieved on Day 8 using Wilcoxon's matched pairs signed-ranked test, it was revealed that the group's mean DT improved significantly between the first (*Mdn* = 7.5) and the last (*Mdn* = 25) day of training, $z = -3.06$, $p = .002$, $r = -.88$.

In terms of both perseverance and training efficiency, therefore, the children who underwent the high intensity training were able to wait until the end of the delay intervals in order to obtain the maximum possible rewards.

Figure 5.8. Mean completion rate (Figure a) and mean Delay Time (Figure b) calculated per block across the second study's 8 training days.
Figure b includes the mean Delay Time achieved by the children in Study 1 for comparison purposes.



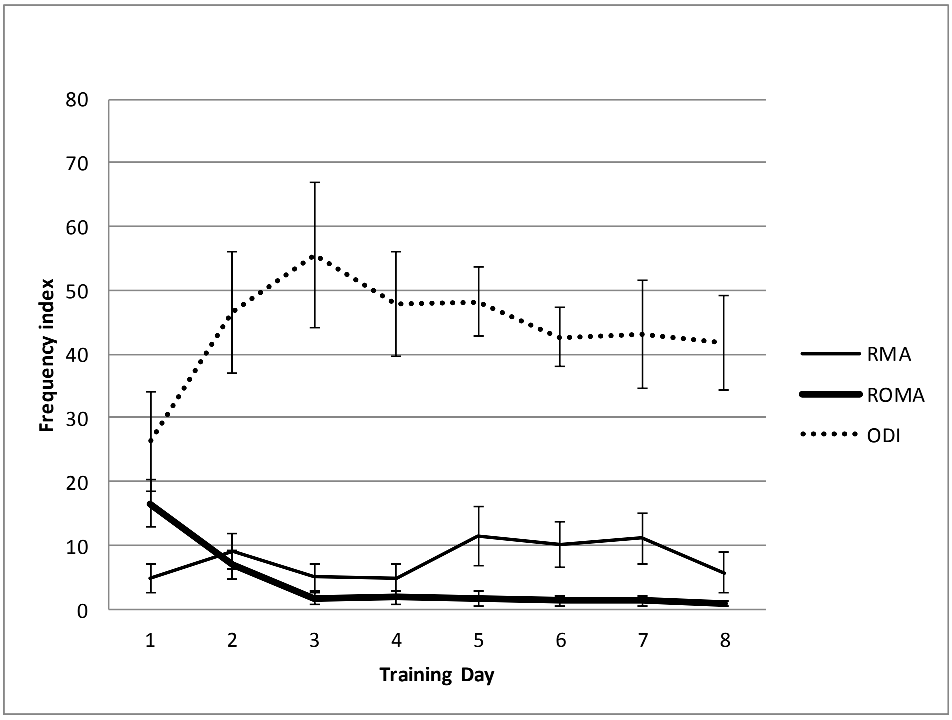


5.4.3.3. Changes in observed waiting behaviour

Child waiting behaviour was coded within the following behavioural categories: Restless motor activity (RMA), Reward-oriented motor activity (ROMA) and the Overall Distraction Index (ODI) aggregate category. As in study 1, an adjusted frequency index was created for each behavioural category in order to correct for the fact delay intervals were becoming longer as children progressed through the training, and as a result more behaviours were likely to occur per delay interval. This was calculated as the total number of times a particular behavioural category was observed divided by the total amount of delay time for which an inhibited or a partially-inhibited score was achieved. The adjusted frequencies per category are presented in Figure 5.9. This figure indicates that as the training progresses children engage in less and less reward-oriented behaviours and in more off-task activities, such as talking to the experimenter (V) or playing with experimental materials (D) while waiting (behaviours included in the ODI category).

Figure 5.9. Observed behaviours per category.

Observed behaviours per category, with all the off-task/distraction categories collapsed into one aggregate category named Overall Distraction Index (ODI): Restless motor activity (RMA); Reward-oriented motor activity (ROMA); the categories making up the ODI aggregate category are: Aimless motor activity (AMA); Looking away (LA); Verbalisation (V) and Out-of seat (OS).



A series of Friedman's ANOVAs were conducted on the adjusted frequencies of observed behaviours. Results indicated that levels of children's restless motor activity (RMA) did not change significantly over the course of the training, $\chi^2(7) = 12.14$, $p = .096$, $r = 1.29$, nor did overall levels of distraction and off-task behaviours as measured by the ODI, $\chi^2(7) = 12.22$, $p = .094$, $r = 1.3$. Changes in children's levels of reward oriented activity across the eight training days were, however, found to be significant, $\chi^2(7) = 24.47$, $p = .001$, $r = 2.61$, but despite the significant overall effect none of the post-hoc pairwise comparisons between training days indicated a significant difference after the probability values had been adjusted for the number of comparisons made.

To assess consistency in coding between the two observers, Cohen's kappa was calculated separately for each behavioural category. Results indicated satisfactory levels of inter-rater agreement with the kappa statistic ranging from .49 to 1, with most categories achieving substantial agreement (values above .61) according to the Landis and Koch interpretation criteria (Landis & Koch, 1977). More specifically, the values obtained per category were as follows: .49 for Restless motor activity (RMA); .67 for Aimless motor activity (AMA); .65 for Reward-oriented motor activity (ROMA); .58 for Looking away (LA); 1.0 for Verbalisation (V); .75 for Distracted Attention (D) and .85 for out-of seat (OS).

A series of Spearman's rho correlations ($\alpha = .05$, two-tailed) were used to explore the relationships between the adjusted frequencies per behavioural category and children's training efficiency on Day 1 (Table 5.8). Confirming the findings of the previous study, the results indicated that training efficiency was negatively correlated with reward-oriented activity (albeit not significantly so in this sample). Importantly, children's overall levels of off-task and distraction activities as measured by the Overall Distraction Index were significantly negatively correlated with reward oriented behaviours even on the first day of training.

Table 5.7. Spearman's rho correlations between the adjusted frequencies per behavioural category and children's mean Delay Time on Day 1

		1	2	3	4
1	mDT-Day 1				
2	Adjusted RMA	-.28			
3	Adjusted ROMA	-.57	.24		
4	Overall Distraction Index	.36	-.54	-.66*	

Note: mDT: mean Delay Time; RMA: Restless motor activity; ROMA: reward-oriented motor activity.

5.4.3.4. Properties of the Delay Aversion Self-report Scale for pre-schoolers (DASS-P)

The internal consistency of the DASS-P scale was low (Cronbach's alpha = .21). A correlational analysis conducted to explore its convergent validity yielded similar results. Children's self-reported delay aversion did not seem to correlate with any of the other outcome measures (Table 5.9).

Table 5.8. Correlations of DASS-P scores with all other pre-/post-training measures at T1

		1	2	3	4	5	6	7	8
1	DASS-P								
2	Teddies	-.07							
3	Cookie Delay	.09	.08						
4	SDQ Total	-.03	-.06	-.13					
5	SDQ Hyperactivity	-.12	-.05	-.13	.89**				
6	QDQ Total	-.11	-.20	-.19	.73**	.86**			
7	BRIEF-P ISCI	-.10	-.17	-.32	.84**	.84**	.76**		
8	BRIEF-P GEC	-.07	-.11	-.30	.94**	.86**	.72**	.90**	

Note: DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

5.4.3.5. Training effects

One-way analyses of variance (ANOVA) with group as the fixed factor and baseline performance as the dependent variable did not show any significant differences between groups for any of the measures at baseline (Table 5.10).

Table 5.9. Demographic and baseline characteristics per group with means (SD) and statistical comparisons

	HI Group (N = 12)	LI Group (N = 13)	Control Group (N = 13)		
	M (SD)	M (SD)	M (SD)	*F*	*p*
Age (yrs)	4.84 (0.28)	5.00 (0.30)	4.94 (0.34)	<1	NS
Teddies	.45 (0.19)	.42 (0.10)	.49 (0.19)	<1	NS
Cookie Delay	.89 (0.25)	.91 (0.11)	.93 (0.12)	<1	NS
DASS-P	2.75 (2.38)	2.80 (2.41)	1.59 (1.80)	1.20	NS
SDQ-Total	3.89 (4.55)	5.23 (5.54)	4.69 (5.22)	<1	NS
SDQ-Hyperactivity	2.13 (1.93)	3.15 (3.34)	2.31 (2.93)	<1	NS
QDQ-Total	11.25 (6.12)	14.15 (9.58)	11.08 (8.55)	<1	NS
BRIEF-P GEC	42.58 (7.09)	43.92 (7.22)	42.85 (7.82)	<1	NS
BRIEF-P ISCI	40.17 (3.64)	43.46 (6.77)	43.08 (9.55)	<1	NS

Note: DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

A series of 3 (group) x 2 (time) repeated measures ANOVAs were conducted on all pre-/post-training outcome measures, with age added as a covariate to the repeated measures design. The descriptive statistics for these analyses are presented in Table 5.11. Significant main effects of age (as a between-subjects factor) emerged for two outcome measures, DASS-P scores, $F(1,34) = 5.73$, $p = .02$, partial $\eta^2 = .14$ and QDQ scores, $F(1,34) = 5.94$, $p = .02$, partial $\eta^2 = .15$, and for this reason the covariate was retained for all analyses. No age by time interaction effects emerged. Summary statistics are presented in Table 5.12.

Table 5.10. Descriptive statistics per group (means and standard deviations).

	HI Group (N = 12)		LI Group (N = 13)		Control Group (N = 13)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Teddies	.45 (0.19)	.48 (0.19)	.42 (0.10)	.45 (0.21)	.49 (0.19)	.43 (0.11)
Cookie Delay	.89 (0.25)	.89 (0.12)	.91 (0.11)	.85 (0.17)	.93 (0.12)	.87 (0.19)
DASS-P	2.75 (2.38)	2.79 (2.33)	2.80 (2.41)	3.01 (1.79)	1.59 (1.80)	2.31 (1.31)
SDQ-Total	3.89 (4.55)	3.83 (4.13)	5.23 (5.54)	3.23 (3.32)	4.69 (5.22)	3.38 (3.73)

	HI Group (N = 12)		LI Group (N = 13)		Control Group (N = 13)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
SDQ-Hyperactivity	2.13 (1.93)	2.33 (2.42)	3.15 (3.34)	2.46 (2.47)	2.31 (2.93)	2.08 (2.81)
QDQ-Total	11.25 (6.12)	10.83 (6.60)	14.15 (9.58)	12.15 (7.10)	11.08 (8.55)	11.31 (6.97)
BRIEF-P GEC	42.58 (7.09)	41.33 (4.87)	43.92 (7.22)	42.92 (4.41)	42.85 (7.82)	41.08 (4.42)
BRIEF-P ISCI	40.17 (3.64)	40.92 (3.73)	43.46 (6.77)	43.38 (5.41)	43.08 (9.55)	41.00 (5.76)

Note: DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

Table 5.11. Summary table of 3 (group) by 2 (time) repeated measures ANOVAs on pre-/post training measures.

	Group			**Time**			**Time by group**			**Time by age**		
	F	p	$_p\eta^2$	F	p	$_p\eta^2$	F	p	$_p\eta^2$	F	p	$_p\eta^2$
Teddies Task	F(2,34)=.25	.78	.01	F(1,34)=.06	.80	.00	F(2,34)=1.21	.31	.46	F(1,34)=.07	.80	.00
Cookie Delay	F(2,24)=.1	.91	.01	F(1,24)=2.3	.14	.06	F(2,34)=.17	.84	.01	F(1,34)=2.6	.12	.07
DASS-P	F(2,34)=1.15	.32	.06	F(1,34)=2.3	.14	.00	F(,34)=.17	.84	.46	F(1,34)=2.6	.12	.01
SDQ Total	F(2,34)=.21	.82	.01	F(1,34)=2.72	.11	.07	F(,34)=2.68	.83	.14	F(1,34)=2.18	.15	.06
QDQ Score	F(2,34)=.84	.44	.05	F(1,34)=1.4	.25	.04	F(,34)=1.38	.27	.07	F(1,34)=1.24	.27	.04
BRIEF-P GEC	F(2,34)=.6	.55	.04	F(1,34)=1.46	.24	.04	F(,34)=.12	.89	.01	F(1,34)=1.2	.28	.03

Notes: DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behaviour Rating Inventory for Executive Function - Preschool Version - Global Executive Composite.

5.4.4. Discussion

The most striking outcome of the second feasibility study was the success of the modifications to the Waiting Game training protocol. Children who were trained on the adapted Waiting Game were able to sustain the delays presented to them successfully, i.e. in a way that maximised their rewards. Children's average delay time significantly improved between the first and the last day of the training, with a large effect size (r = .88) obtained for this comparison. Similarly, the majority of children were able to persevere with the training as evidenced by the high trial completion rate and the low dropout rate recorded in this study. This was a dramatic difference to the extensive non-compliance to protocol of the previous study.

The second objective of the current study was to investigate the reliability of the Observational Checklist and also to confirm the observational analysis findings of the first feasibility study. The inter-rater reliability of the WG's observational coding system was satisfactory (with the kappa statistic ranging from .49 to 1). With regard to changes in observed behaviours over the course of the training, children were observed to engage in significantly less reward-oriented behaviours and in more off-task and self-distraction activities as the training progressed. In fact, these two types of activities (reward-focused as opposed to distraction-focused) were found to be significantly negatively correlated even on the first day of training.

Interestingly, these findings are not consistent with of our initial hypotheses in relation to waiting behaviour during training: children were found to be engaging in more and more inattentive behaviours as the training went on, and not less as it had originally been postulated. In addition, children's activity levels were largely unaffected over time, in the face of the earlier prediction that training might result in reduced levels of over-activity. These changes in behaviour albeit not directly anticipated by the delay aversion model, are still in line with its basic tenet that inescapable delays may lead to children engaging in off task behaviours that help "speed up" the passage of time and ultimately alter the subjective experience of delay (Sonuga-Barke, 2003). In this context, children may not have become less delay averse during training, as evidenced by their increasing levels off task behaviours, but they still showed increases in waiting efficiency, as evidenced by the fact they were able to sustain increasing delays successfully across the training sessions.

This pattern of behaviour is also consistent with the classic literature on the effects of attention to rewards in a delay of gratification context (Mischel & Ebbesen, 1970; Mischel, et al., 1972; Peake, et al., 2002). This line of research has established that an attentional focus on rewards undermines children's ability to delay gratification. Further, it has also demonstrated that compared to older children four-year olds are very ineffective in directing their attention away from the rewards at hand, preferring instead to look at the rewards and to think about them, thus compromising their ability to wait (Mischel, et al., 1989). Our findings support the former assertion that orienting attention to rewards correlates with unsuccessful waiting. We have also shown that even children as young as four spontaneously develop self-distraction strategies that facilitate effective waiting over time without having been explicitly instructed to do so. This finding suggests that children have the ability to self-regulate when given the opportunity to practise waiting, without having been taught explicit strategies, which is one of the fundamental assumptions of the WG training paradigm.

The third aim of this feasibility study was to investigate the differential effects of training on two versions of the training task, a high intensity one and a low intensity one. The results indicated that the gains in performance on non-trained tasks and in teacher-rated behaviour pre and post training were minimal and did not point to any differences between the three groups. This pattern of results could be attributed to a number of factors: firstly, the intensity of the training may have been too easy for the children, even the ones training in the high intensity condition. This could potentially explain why children who trained on either condition exhibited performance virtually the same as children who did not receive any training at all. Another possibility is that the two measurement points (T1 and T2) were so close to one another no detectable behavioural changes could have possibly occurred, especially in terms of global behaviours captured by the three behavioural rating scales. Lastly, the possibility should always be entertained that behavioural changes incurred in the context of such a structured learning/training paradigm may not readily generalise to other related constructs particularly after such a brief period of training.

5.4.4.1. Limitations

Because of its very nature, the current feasibility study has important limitations. Firstly, as was the case with the first study, the fact this was a small-sized pilot study has meant it lacked the statistical power necessary to investigate the reported results robustly and definitively. Secondly, it needs to be acknowledged that the researcher who implemented the intervention also conducted the pre-post-training assessments, which meant she was not blind to group allocation (the teachers, however, who were the main informants of the study were). Thirdly, both feasibility studies used a normative sample to test aspects of the intervention designed to improve the delay tolerance of children with delay related difficulties. This was deemed appropriate as the work described in this chapter was novel and there were many aspects to the training protocol which were appropriate to test with a normative sample (such as the psychometric properties of the observational system, the motivational value of different rewards in this age group, etc.). There are, however, different aspects of the WG training task that are difficult to evaluate on the basis of the normative data presented in this chapter, such as the question of what constitutes an appropriate level of intensity of training for a clinical sample.

5.4.4.2. Future Directions: issues to address

Intensity of training. It is very difficult to tell which of the modifications to the protocol brought about the most change in children's behaviour but anecdotal evidence suggested that children were motivated to wait until the end of each delay interval because they were rewarded for it. To avoid running the risk of setting the intensity of training too low, it is proposed that in future implementations of the WG task the additional reward contingency is retained, but the training schedule is reverted back to that used in Study 1 (but using the first training block only). This schedule will use a smaller number of trials to implement larger delays in a more compressed time scale (5 training days). It will also avoid having a break in the middle of the session, which was thought to contribute to children's fatigue and high drop out rate.

Selection of control group. Because the findings of the current study were unfortunately not informative in terms of the suitability of the low intensity WG task as a comparison task, it is proposed that a waiting list group would be a more pragmatic choice for the first evaluation trial.

Selection of sample. As mentioned previously, children participating in the next study will be selected on the basis of their delay related difficulties, to facilitate conclusions relating to a clinical sample.

Chapter 6: A Small Scale Randomized Controlled Trail of the Effects of Delay Training ("The Waiting Game") on Young Impulsive Children

6.1. Aims of the chapter

The main aim of this chapter is to report on the findings of a small scale randomised controlled trial (RCT) designed to evaluate the training efficiency and the training effects associated with the adapted WG training task in a sample of preschool-aged children presenting with high levels of delay related difficulties. This chapter will also assess the feasibility of the training programme in terms of its acceptability, retention rate and effect sizes in this sample with a view to facilitating conclusions as to whether a full scale evaluation of the programme should be undertaken using a clinical ADHD sample.

6.2. Introduction

Chapter Five presented the process of developing a delay shaping training procedure (The Waiting Game; WG) and tested its feasibility in two samples of typically developing children. The aim of the current study was to evaluate the WG task when implemented as an intervention for children at risk for ADHD. More specifically, and in line with the early intervention approach to theory driven and pathway-specific treatment development outlined in Chapter Two, the focus of the current study was children selected for their difficulties with waiting and who were at risk for developing motivationally based ADHD type problems. Within this framework, the current trial provides an exploratory evaluation of the WG training programme in a high-risk sample allowing inferences relating to a clinical ADHD sample to be drawn.

The time of school entry was selected as a suitable intervention point. Children start to develop the ability to wait as they are moving away from toddlerhood and into preschool (Calkins, 2007). In formal schooling, children are largely expected to be able to control their drives for immediate outcomes (such playing or focusing on activities of their preference) and to focus instead on teacher-led activities for longer periods of time (Blair, 2002; Graziano, Reavis, Keane, & Calkins, 2007). Understandably, many children face difficulties with this skill and especially children at risk for developing behaviour problems (Supplee, Shaw, Hailstones, & Hartman, 2004).

In the United Kingdom, children typically begin school in the reception year of primary school at the start of the academic year following their fourth birthday. The Year R curriculum is guided by the Early Years Foundation Stage (EYFS) statutory framework which sets the standards for the learning, development and care of children from birth to five years of age in all types of Early Years settings. The transition from nursery and pre-school settings to Year R often marks a change to a longer and more structured school day, where children are expected to practice waiting more frequently and for longer periods. This study set out to recruit a sample of children facing particular difficulties waiting during delays in this period of transition. We screened all children starting the Year R in five schools across Hampshire in order to select children who were identified by their teachers as having elevated levels of delay related difficulties compared to their peers.

This recruitment strategy also determined the context of our intervention. Given teachers' involvement during the screening and the recruitment process of this study, it was deemed appropriate for the intervention to be delivered in the school setting and for teachers to be the study's main informants. Eventually the WG training task would ideally be rolled out as one component of a more complex early intervention programme targeting preschool children at risk for ADHD. Such preventive programmes should be accessible to large numbers of children and ideally designed to be helpful to all children regardless of at-risk status, thus making them ideal to implement in school or community based settings (Halperin, Bedard, et al., 2012). Parental involvement is also key as the ultimate goal would be for these therapeutic activities to be fun, engaging and easy to incorporate into daily life (Halperin, Bedard, et al., 2012) and for the intervention to be delivered via a collaborative school-home format (see Pfiffner, Villodas, Kaiser, Rooney, & McBurnett, 2013 for an example). However, at this stage in the development of the WG task and because of the exploratory nature of its current implementation, parental collaboration was not considered essential for this study.

Full details on the methodology and the findings of the current trial are reported in the remainder of this chapter following the CONSORT (Consolidated Standards of Reporting Trials) statement, as extended to trials of non-pharmacologic treatments

(Boutron, Moher, Altman, Schulz, Ravaud, & Group, 2008; Boutron, Moher, Altman, Schulz, Ravaud, & Grp, 2008). The CONSORT statement is a well-established standard for reporting randomised controlled trials. It consists of a 22-item checklist and a flow diagram designed to improve the quality of RCT reporting and to aid with the critical appraisal of the validity and applicability of trial results. The statement's extension to non-pharmacologic treatments represents an effort to address specific issues that apply to trials of non-pharmacologic interventions, which include behavioural therapy and psychological interventions among others. The extended CONSORT checklist provides the structure of this chapter and can be found in Appendix A.8.

6.3. Methods

6.3.1. Study design and procedures

All children were recruited though their schools via the Research in Partnership Scheme organised and run by the academic unit of Psychology, University of Southampton. A form containing detailed information about the research was sent out to schools participating in the scheme inviting them to take part. Schools that indicated an interest in the project were subsequently visited by the researcher to discuss all the practicalities involved in the trial. Five primary and infant schools agreed to take part (two schools from the Southampton City area and the other three located in non-metropolitan surrounding districts). Informed consent was obtained from all participants (parents and teachers) prior to initiating proceedings. All procedures were approved prior to the start of the study by the Ethics Committee for the Faculty of Social and Human Sciences of the University of Southampton (please see Appendix A.9 for all relevant documentation addressed to schools and Appendix A.10 for the information sheets and consent forms addressed to parents). Schools received book vouchers worth £25 for every classroom that took part in the study as compensation.

During the screening phase of the study, classroom teachers were asked to complete the QDQ rating scale for all children in their classrooms. Subsequently, teachers were asked to complete two additional rating scales which assessed the behaviour of the children who were selected to participate in the training phase of the study, as part of the pre- and post-training assessments. Children who were randomised into one of the two experimental groups also underwent a neuropsychological assessment s as part of their pre- and post-training tests. All sessions, including the two assessment and the five training sessions, were held on separate school days over the course of three weeks and took place in a quiet room in the children's school. Assessment sessions lasted for approximately 20-25 minutes and training sessions (n = 5) about 10 minutes each. Children were offered breaks during the assessment sessions but because the training sessions were so brief they were not offered breaks during these. Children allocated to the waiting list group did not receive any training sessions.

6.3.2. Participants

A total of 209 children were screened using the QDQ behaviour rating scale. Children whose QDQ scores were in the top 20 percent of their class indicating high delay difficulties met the main eligibility criterion for this trial. Exclusion criteria included a statement of special educational needs or previous diagnosis of a pervasive developmental disorder or in the case of children with English as an additional language, poor comprehension of spoken English as assessed by their teachers. Forty seven children were selected with high QDQ scores. Ten of those children were excluded either because they met the study's exclusion criteria (n = 6) or because of other practical considerations raised by their teachers (n = 4, e.g. some children were not yet attending full school days or would not be available during the data collection period because of other commitments, e.g. rehearsals for school plays)(see also Figure 6.1). Eligible children for whom parental consent had been obtained were randomised into two conditions: a training condition and a waiting list group. In total 37 children were randomised, with 18 assigned to the training group and 19 to the waiting list group. Their ages ranged from 4.2 to 6.1 years (M = 4.8, SD =.47). For a detailed description of the randomisation method please see section 6.3.7. Data collection for this study took place from October 2012 to April 2013.

6.3.3. Interventions

Children allocated to the training condition trained on the adapted WG training task, as described in section 4.4.2.2, but following a more demanding training schedule. In accordance with the conclusions reached after the implementation of less a challenging schedule in the second feasibility study, the training protocol used in the current study was more intensive. The current schedule used a smaller number of trials to implement larger delays in a more compressed time scale (over 5 training days). In addition, it introduced two additional training levels to achieve more precise tailoring to children's abilities. Furthermore, a free waiting trial was added at the end of every training day. For this, children were asked to wait for as long as they could manage for a self-selected reward. All aspects of the WG training task and protocol, including instructions on how to administer and score the task, had been fully manualised (please see Appendix A.11 for the complete WG manual). The inter-rater reliability of the task's scoring and observational systems were examined and reported in Chapter Four. To ascertain inter-rater agreement in a sample of impulsive children, a second independent rater observed six children on their last (fifth) day of training. That rater received the full scoring instructions described in the adapted WG manual and no further training.

6.3.4. Objectives and hypotheses

The current trial had three main aims: firstly, to evaluate the adapted WG task both in terms of training efficiency and in terms of children's perseverance to complete the training in a sample of children at risk for ADHD; secondly, to establish the inter-rater reliability of the adapted Observational Checklist in this sample and use observational data to confirm that increases in delay tolerance are accompanied by measurable changes in waiting behaviour and in particular by a shift of focus from the rewards at hand to off-task behaviours, in line with the findings of the previous studies; thirdly, to investigate improvements in delay tolerance after training on the adapted WG task as measured both by children's performance on untrained delay tasks and by teacher ratings of delay related behaviours observed in the classroom setting.

It was hypothesised that gains in delay tolerance after training on the adapted WG task would be higher for the experimental group, compared with the waiting-list group. It was also hypothesised that the adapted WG training protocol would lead to improved completion rates and gains in delay times across the training period. Finally, it was predicted that children would engage in less reward-oriented behaviours while waiting to focus on other distractions, as suggested by the findings of the two feasibility studies described in Chapter Four.

6.3.5. Outcomes

6.3.5.1. Neuropsychological measures

The Teddies task. As described in section 4.3.1.2. This was the primary outcome measure of this study in order to maintain continuity with the previous studies and between the power analysis and data analysis methods.

The Bee task. As described in section 4.3.

6.3.5.2. Behavioural Rating scales

The revised Delay Aversion Self-report Scale for pre-schoolers (DASS-P). This was a self-report measure of delay tolerance, developed specifically for this research. It was first introduced in the second feasibility study (in section 5.4.2.4) but was revised as its psychometric properties were not found to be satisfactory (section 5.4.3.4). Two changes were introduced in this version: firstly, its response format was simplified. Children were now given three response options for each question (as opposed to five) because it was observed that children's responses tended to gather around the two extremes of the five picture rating scale (in other words, very few children picked any of the three middle response options). To address this issue, the scale was reduced to a three point scale and all three response options were clearly worded, e.g. the three response options to the question "Do you like waiting for things?" were clearly stated as "Yes/No or I don't know"). The second change introduced involved the inclusion of an additional item in an attempt to improve the scale's poor internal consistency. In this version of the scale, responses to every item ranged from 0 to 2, thus making the total DASS-P score range from 0 to 6, with higher scores indicating higher aversion to delay. Because this scale was used to monitor children's delay aversion during the training and as part of the pre-and post-training assessments, two slightly different versions of the questions were devised: a trait version, inquiring about children's feelings towards waiting in general (which was used during the pre- and the post-test) and a state version, asking children about the waiting they had to do while training (this version was used at the end of every training day). Full details on the scale's items and response options can be found in Appendices A.12 and A.13)

Teacher questionnaires. As in section 5.4.2.4.

6.3.6. Sample size

Sample size calculations were computed using an internet-based programme (GLIMMPSE, version 2.1) specifically developed to deal with sample size and power calculations for repeated measures and longitudinal designs (Guo, Logan, Glueck, & Muller, 2013). To estimate the required sample size, this programme needs input on the desired power value, the Type I error rate, the target hypothesis being tested, the difference in the pattern of means for which power is being sought, the variances of the response variables and the correlations among the response variables (Guo, et al., 2013). As per standard practice, the desired power value was set at 0.8 and the Type I error rate at 0.05. The last three items in the list were estimated on the basis of data from the second feasibility study. More specifically: the observed mean values of the primary outcome measure (The Teddies Task) for the high intensity and the control groups were used in the model; equal variances between the repeated measurements were assumed because measurements were taken so close in time; regarding correlations, again the observed correlations between time points and groups were inputted in the model (two-tailed Pearson correlations $r = 0.6$ and $r = 0.4$, respectively). In terms of hypothesis testing, GLIMMPSE can test for an intervention (group) main effect and for a time × intervention interaction effect. A model was run to calculate a sample size that would allow detection of a time x intervention interaction effect (i.e. testing whether the trend of change over time is different between the intervention and the non-intervention group). The sample size specified as necessary to detect a time x intervention effect with 80% power and at a = 0.05 was 41. These calculations indicate that if the current trial manages to recruit its target total sample size of 40 participants, it will be powered to detect a time x group interaction effect.

6.3.7. Randomisation

Randomisation was done in blocks (of two and four) to avoid disparities in the sizes of the two groups. The generation of the allocation sequences was done using the website www.randomiser.org which uses a JavaScript random number generator to produce customized sets of random numbers. To ensure that classroom teachers, who were the main informants of the study, could be kept blind to group allocation, arrangements as to which children would be called out of the classroom and when were made with classroom assistants whenever possible.

6.3.8. Statistical methods

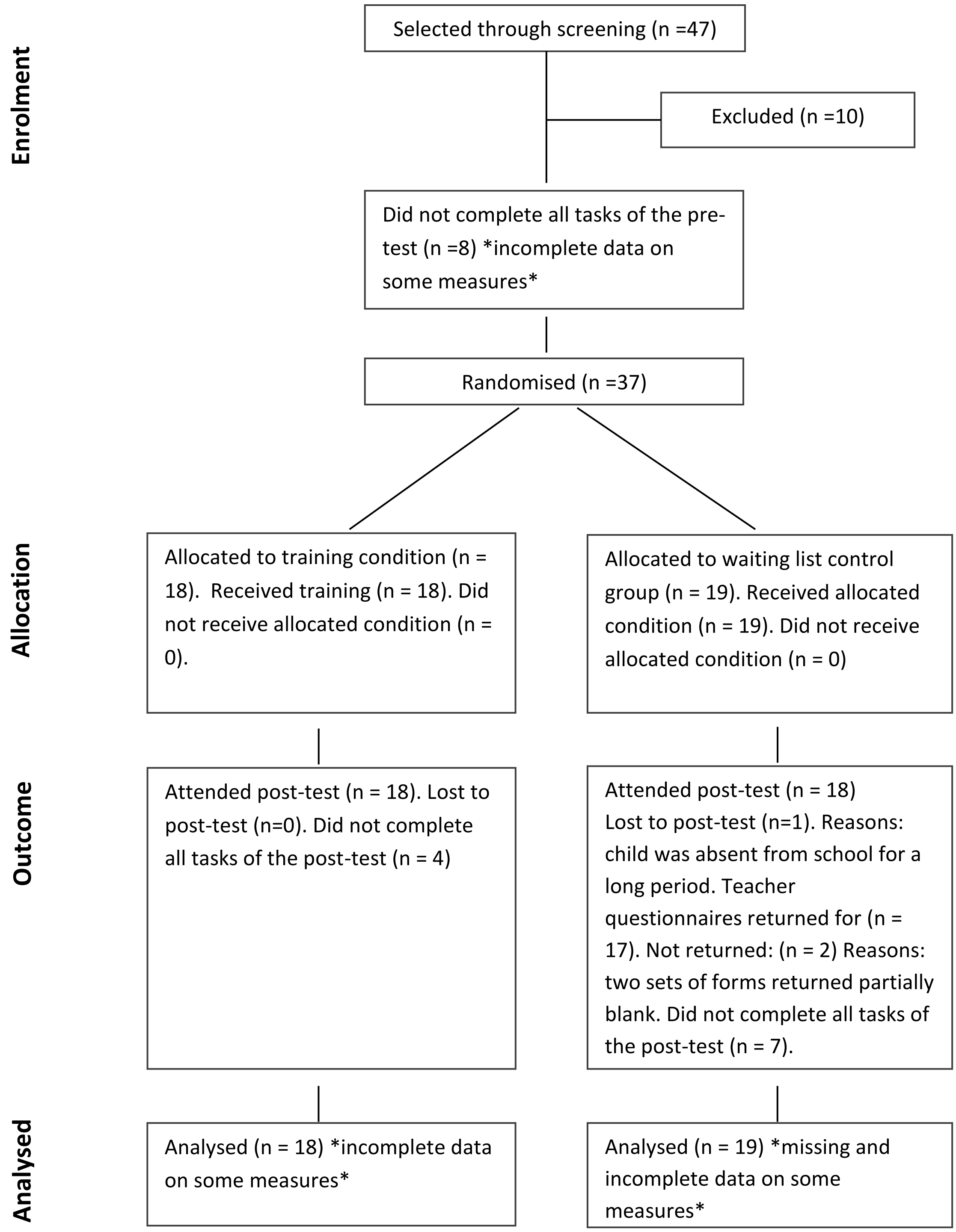
Baseline Characteristics: descriptive statistics were used to describe the demographic and baseline clinical characteristics of the two groups. To check for differences between the two groups, independent samples t tests were used for continuous variables and the chi square test for categorical variables. Training efficiency: descriptive statistics were also used to assess children's training efficiency and perseverance with the training. To further assess the training efficiency of the training group, the mean delay time and the mean free waiting time achieved on the first day of training was compared to those achieved on the last day, using Wilcoxon matched pairs signed-ranked tests (see also section 6.4.1. for a justification of the use of non-parametric procedures). Descriptive statistics and Friedman's ANOVAs were used to investigate changes in waiting behaviour over the course of the training as measured by the Observational checklist. To assess consistency in coding between two independent observers, a series of Cohen's kappas was calculated, separately for each behavioural category. A bivariate correlational analysis was used to explore the relationships between observed behaviours and children's training efficiency on the first day of training. Measure development: Cronbach's alpha and a series of Pearson's correlations were conducted to explore the internal consistency, test-retest reliability and the convergent validity of the revised DASS-P scale. Training Effects: a series of 2 (group) x 2 (time) repeated measures ANOVAs were conducted on all pre-/post-training outcome measures, with age and sex added as a covariates to the repeated measures design, to investigate between-group and interaction effects on children's test-retest performance. Because of the small size of the sample and the fact one therapist administered the intervention across the whole sample, no cluster analyses were conducted on the data. Effect sizes are reported using partial eta squared for the repeated measures ANOVAs and the effect size estimate r for the non-parametric procedures. Effect sizes are interpreted as .1, .3 and .5 reflecting small, medium and large effect sizes respectively (Cohen, 1992). The significance level (α) was adjusted for multiple testing using the Bonferroni correction.

6.4. Results

6.4.1. Participant flow and numbers analysed

As described in detail in section 6.3.2, 209 children were screened for this trial, 47 were selected through screening and 37 were randomised into groups (see also Figure 6.1). In terms of missing data, the drop out rate from the study was very low, with no children dropping out of the training phase of the study (n = 0) and only one child being partially lost to the post test (the child was absent from the T2 assessment, but his teachers provided complete T2 questionnaire data). For two other children, teachers failed to return complete T2 questionnaires. When a missing data analysis was conducted, Little's MCAR test indicated that data was missing at random (χ^2 = 3.23, *df* = 57; *p* = 1.0). Following this result and considering the amount of missing data was less than 1%, the Expectation Maximization (EM) algorithm was used to compute maximum likelihood estimates of the missing values (Jelicic, et al., 2009). There was however a separate issue with incomplete data on the two neuropsychological delay tasks (the Teddies Task and the Bee Task). As had happened in the first feasibility study, children on occasion asked to abandon the delay tasks early, before they had completed all available trials. In the current trial this very rarely happened during the training, with only two instances of early termination recorded on the last day of training. But a large number of children (8 children at T1 and 11 children at T2) abandoned either one or both of the computerised delay tasks administered to them during the pre-post training assessments.

Figure 6.1. Consort Diagram



To avoid making data imputations in these cases, children's performance was estimated on the basis of the number of trials they had completed. Because both delay tasks calculate delay tolerance as a function (percentage) of the overall number of trials completed, a slight modification of the dependent variables was necessary: for the Teddies task, instead of calculating the percentage of LL responses, a composite "Teddies score" was computed with 1 point awarded for every SS response and two points for every LL response, up to the point of task completion. This composite delay index allowed an estimation of waiting efficiency (as LL responses have double the weight of SS responses) while still taking into account task perseverance -given that a child who terminates the task early is likely to achieve a low Teddies score regardless of response preference. Following the same reasoning, the total amount of delay recorded across all completed trials was used as the main index of the Bee task. Again, this was an index of children's waiting efficiency while performing that task but at the same time sensitive to early termination (given that early termination would be expected to result in fewer trials and in turn to a smaller amount of overall delay tolerated across the whole task).

The proposed modifications to the delay variables described above allowed the preservation of the current trial's original sample size. To reiterate, incomplete data from the two computerised delay tasks were retained for analysis on the grounds that early exit was seen as a delay escaping tactic and that the point of task termination should contribute towards the calculation of the overall delay tolerance scores. However, because this approach was untested and the proportion of incomplete data was high (varying from 19% to 25%), all analyses on delay variables were repeated using complete cases only (and using the original dependent variables, i.e. percentage of LL response in the case of the Teddies Task and mean delay per trial for the Bee Delay task. Results were thus reported for both sets of variables where appropriate: Teddies Mod and Bee Mean Delay Mod (modified variables) and Teddies and Bee Mean Delay Comp (complete cases only).

6.4.2. Data processing and screening

Prior to analysis all data were examined using IBM SPPS Statistics 20 for accuracy of data entry, missing values and fit between their distributions and the assumptions of the statistical analyses used in this section, according to data screening procedures described by Tabachnick and Fidell (2007).

Screening procedures for outliers and checks for the normality of variables were carried out separately for the two groups according to the protocol for screening grouped data outlined by Tabachnick and Fidell (2007). The entire dataset was scanned for cases with standardised scores in excess of 3.29 ($p < .001$, two-tailed) and six outliers were identified using this method. Five outlying values were located in observation variables and one value was in the free waiting variable. The impact of these outliers was reduced by replacing the outlying values with scores one unit larger or smaller than the next most extreme score in the distribution in line with procedures outlined by Tabachnick and Fidell (2007). Normality checks were run before and after the substitution of outliers. The majority of the training efficiency and observation variables (about two thirds of these variables) were found to be non-normally distributed when tested statistically, and this pattern of results was completely unaffected by the substitution of outliers. Closer inspection of frequency histograms and normal probability plots revealed that the majority of these distributions departed substantially from the assumptions of zero skewness and kurtosis. As in previous studies, this was attributed to the nature of the data sampling methods used, with very high prevalences of tied values (in the case of training variables) and of zero values (marking the non-occurrence of target behaviours during observation and the absence of delay aversion as measured by DASS-P scores) making delay training variables very prone to normality violations. For these reasons, non-parametric procedures were selected to analyse all training and observational data.

A small number of pre-/ post-training outcome variables were also found to deviate significantly from normality when assessed within groups (about one third of Shapiro-Wilk tests reached statistical significance in this group of variables indicating deviations from normality), but given the robustness of the analysis of variance procedure (Wilcox, 2001) the decision was made to proceed with parametric analyses in this case, using repeated measures ANOVAs to evaluate pre-post training effects.

6.4.3. Baseline Data

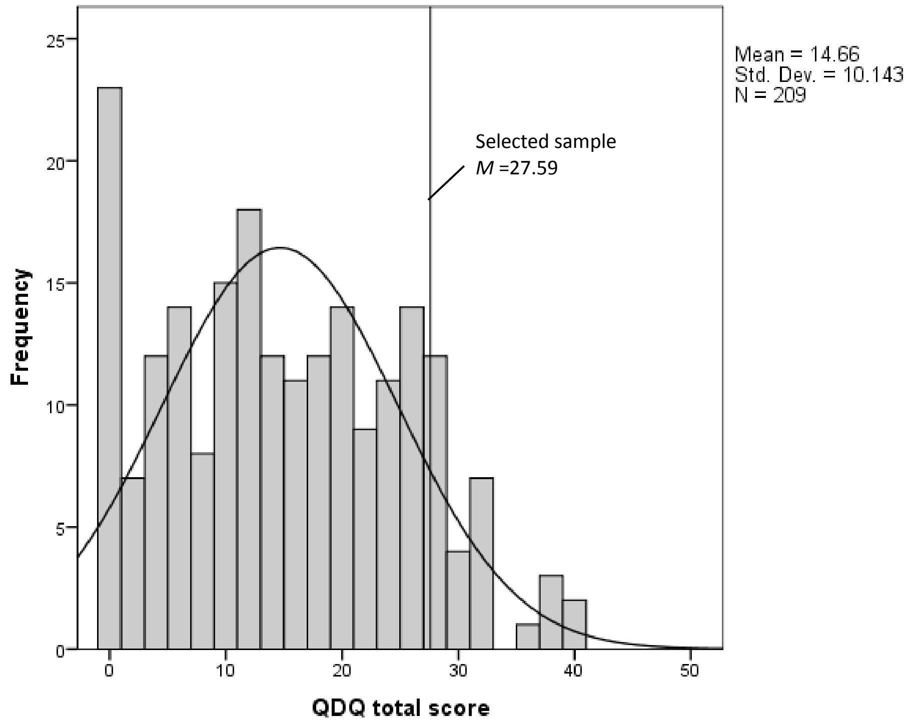
Demographic and baseline characteristics of the two groups are presented in Table 6.1. Sixty three percent of the children met the clinical cut-off for pervasive hyperactivity on the SDQ scale (a score of ≥ 6 on the hyperactivity subscale)(Sayal, Taylor, Beecham, & Byrne, 2002) and 42% met the cut off for elevated BRIEF-P scores with clinical implications (a Global Executive Composite T score of 65 or above)(Sherman & Brooks, 2010). Their mean QDQ score was nearly 1.5 standard deviations above the mean recorded for the screened sample (see Figure 6.2 for the frequency distribution of QDQ scores in the screened sample). The distribution was positively skewed with an accumulation of cases towards the left side of the curve. This distribution shape is indicative of a floor effect and not unusual for a problem-focused measure). No differences were reported between groups at baseline, using independent samples t tests for continuous variables and the chi square test for categorical variables.

Table 6.1. Demographic and baseline characteristics per group with means (SD) and statistical comparisons

	Both Groups (N = 37)	**Training Group** (N = 18)	**Control Group** (N = 19)	
	M (SD)[N]	M (SD)[N]	M (SD)[N]	statistical comparison
Boys (% boys)	24 (65)	13 (72)	11 (68)	$\chi^2(1) = 8.33$, n.s.
Age (yrs)	4.8 (.47)	4.8 (.55)	4.8 (.37)	$t(35) = .06$, n.s.
SQD Hyperactivity ≥ 6 (%)	24 (65)	12 (67)	12 (63)	$\chi^2(1) = .05$, n.s.
BRIEF-P GEC T score ≥ 65 (%)	16 (43)	6 (33)	10 (53)	$\chi^2(1) = 2.17$, n.s.
Teddies Score Mod	13.38 (3.93)	13.68 (3.96)	13.06 (3.98)	$t(35) = .48$, n.s.
Teddies Comp	.49 (.15) [30]	.49 (.18) [17]	.50 (.12) [13]	$t(28) = -.21$, n.s.
Bee Delay Mod	306.98 (172.52)	316.81 (194.03)	296.59 (151.48)	$t(35) = .69$, n.s.
Bee Mean Delay Comp	36.47 (14.70)[29]	34.84 (17.91)[17]	38.77 (8.57)[12]	$t(27) = .70$, n.s.
DASS-P-T	1.08 (1.74)	0.79(1.47)	1.39 (1.97)	$t(35) = .17$, n.s.
SDQ-Total	13.29 (4.94)	12.76 (5.12)	13.85 (4.82)	$t(35 = .90$, n.s.
QDQ-Total	27.59 (7.62)	27.89 (7.94)	27.28 (7.49)	$t(35) = .99$, n.s.
BRIEF-P GEC	62.32 (12.98)	59.68 (11.93)	65.11 (13.78)	$t(35) = .28$, n.s.
BRIEF-P ISCI	44.68 (11.93)	42.11 (10.13)	47.39 (13.32)	$t(35) = .22$, n.s.

Note: SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version-Global Executive Composite; DASS-P-T: Delay aversion self-report scale for preschoolers (trait version); QDQ: Quick Delay Questionnaire-Preschool version; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index; *t* = independent samples t-test; χ^2 = chi square test; n.s. = not significant at α = 0.05 level

Figure 6.2. Frequency distribution of QDQ scores across the screened sample (N = 209)
The reference line drawn on the x axis marks the selected sample mean of 27.59.



6.4.4. Training efficiency

The mean duration of all delay intervals in seconds (mean Delay Time) for which a fully- or partially-inhibited score was achieved was calculated per child per training day. In cases of incomplete training sessions, mean DTs were averaged across the number of trials that had actually been completed. This procedure calculated an index of training efficiency on the basis of children's performance while they trained. The group's mean completion rate (number of trials completed over the total number of trials offered per session) was also calculated per day to provide more information in relation to the children's perseverance with the training. As can be seen from Table 6.2, children's perseverance with the training was high and the number of incomplete trials reported was negligible. Table 6.2 also provides information on the average time children chose to wait in the free waiting trial offered to them at the end of each training day and on children's self-reported delay aversion scores (for a more detailed discussion of the DASS-P scale please also see section 6.4.6).

Table 6.2. Means and standard deviations for all training efficiency measures across the five training days

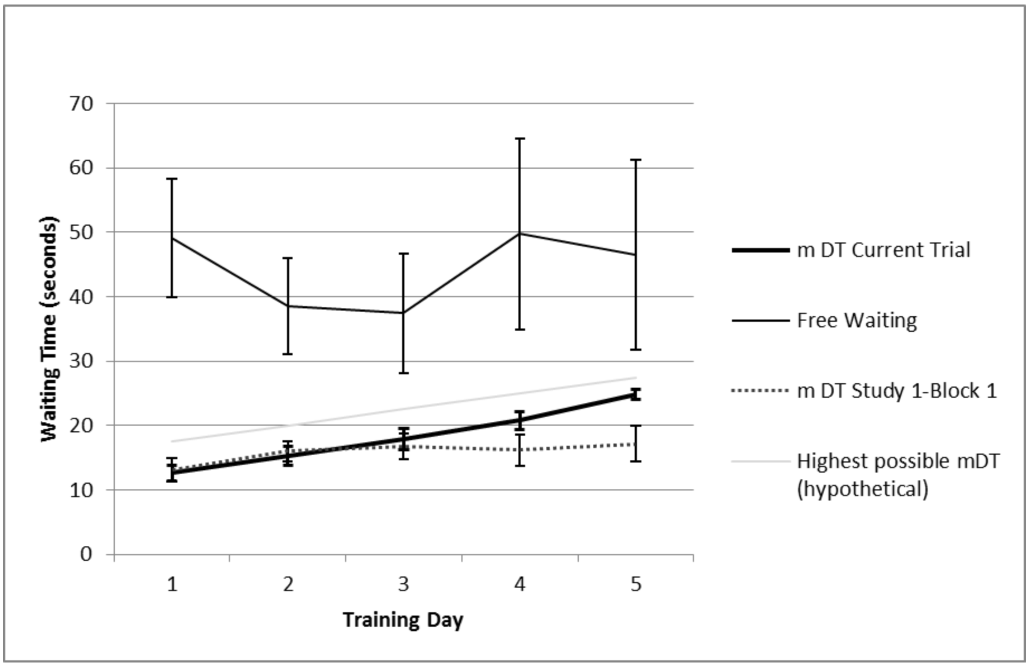
	Day 1	Day 2	Day 3	Day 4	Day 5
Mean Delay Time	12.63 (5.21)	15.33 (6.43)	17.89 (7.25)	20.79 (5.86)	24.89 (3.29)
Completion Rate	1.00 (.00)	1.00 (.00)	1.00 (.00)	1.00 (.00)	.97 (.10)
Free Waiting	49.05 (40.04)	38.53 (32.41)	37.47 (40.42)	49.74 (64.89)	46.56 (62.52)
DASS-P-S (State version)	1.37 (1.77)	1.84 (2.03)	2.00 (2.05)	1.58 (1.43)	1.95 (1.87)

Note: DASS-P-S: Delay aversion self-report scale for preschoolers-state version.

To further probe changes in performance over time, the mean DT achieved on Day 1 was compared with that achieved on Day 5 using Wilcoxon's signed ranked tests. Please also see Figure 6.3 for an illustration of how children's waiting times changed over time. Results indicated the group's mean DT improved significantly between the first (*Mdn* = 13.75) and the last (*Mdn* = 27.50) day of training, $z = 3.84$, $p = .0$ and a large effect size was detected $r = .88$. This result is significant even after the Bonferroni adjustment to the overall alpha level (with α having been lowered to 0.017) has been taken into account. Two further Wilcoxon's tests confirmed that the median duration of the Free Waiting trial recorded on Day 1 (*Mdn* = 32 seconds) was not significantly different from the median Free Waiting trial duration of Day 5 (*Mdn* = 31 seconds), $z = -1.07$, $p = .29$, $r = -.24$; similarly, no significant differences were detected when the median DASS-P-S score obtained on Day 1 (*Mdn* = 1) was compared to that obtained on Day 5 (M*dn* = 2), $z = 1.02$, $p = .31$, $r = .23$.

Figure 6.3. Mean Delay Time (mDT) and average free waiting time calculated across the five days of training.

This figure also computes the highest possible mDT and the mDT achieved by the children in the first block of the Study 1 for comparison purposes.



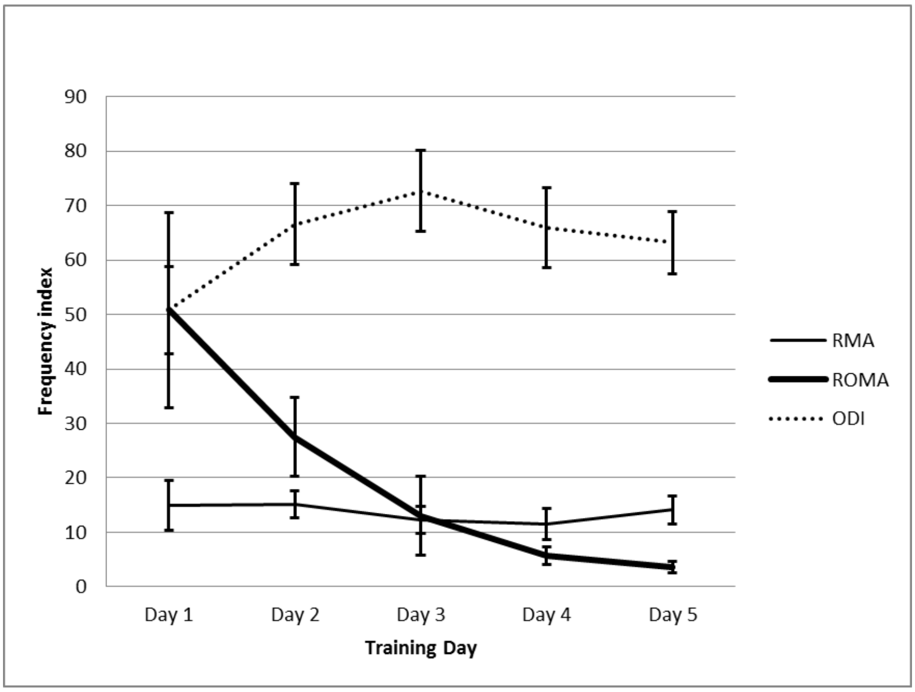
6.4.5. Changes in observed waiting behaviour

As described in Appendix 6.3, child waiting behaviour was coded within the following behavioural categories, specified by the revised Observation Checklist: Restless motor activity (RMA), Reward-oriented motor activity (ROMA) and the aggregate Overall Distraction Index (ODI). As in the previous studies, an adjusted frequency index was created for each behavioural category in order to correct for the fact delay intervals were becoming longer as children progressed through the training, and as a result more behaviours were likely to occur per delay interval. This was calculated as the total number of times a particular behavioural category was observed divided by the total amount of delay time for which an inhibited or a partially-inhibited score was achieved. The adjusted frequencies per category are presented in Figure 6.4. This figure shows that as the training progresses children engage in less and less reward-oriented behaviours and in more off-task activities, such as talking to the experimenter (V) or playing with experimental materials (D) while waiting (both Verbalisation (V) and Distracted attention (D) are behavioural categories included in the ODI category).

Children's restless motor activity levels did not seem to change drastically over time, in comparison.

Figure 6.4. Observed behaviours per category

Observed behaviours per category with all the off-task/distraction categories collapsed into one aggregate category named Overall Distraction Index (ODI): Restless motor activity (RMA); Reward-oriented motor activity (ROMA); the categories making up the ODI aggregate category are: Aimless motor activity (AMA); Looking away (LA); Verbalisation (V) and Out-of seat (OS). Error bars represent the standard error of the mean.



To confirm these patterns of change over time, a series of Friedman's ANOVAs were conducted on the adjusted frequencies of observed behaviours. Results indicated that levels of children's restless motor activity (RMA) did not change significantly over the course of the training, $\chi^2(4) = .65$, $p = .96$, $r = .07$, nor did their overall levels of distraction and off-task behaviours as measured by the ODI, $\chi^2(7) = 5.36$, $p = .25$, $r = .55$. Changes in children's levels of reward oriented activity across the five training days were, however, found to be significant, $\chi^2(7) = 9.68$, $p = .046$, $r = 0.99$, but despite the significant overall effect none of the post-hoc pairwise comparisons between training days indicated a significant difference after the probability values had been adjusted for the number of comparisons made. (Note that if the Bonferroni adjustment was applied to this group of tests, then the overall alpha level would have to be corrected to $\alpha = .017$ and the null hypothesis should be retained for all the tests reported above).

To assess consistency in coding between two independent observers, Cohen's kappa was calculated separately for each behavioural category based on the observation of six children during their last day of training. Results indicated satisfactory levels of inter-rater agreement with the kappa statistic ranging from .59 to 1, with most categories achieving substantial agreement (values above .61) according to the Landis and Koch interpretation criteria (Landis & Koch, 1977). More specifically, the values obtained per individual category as are follows: .58 for Restless motor activity (RMA); 1.0 for Reward-oriented motor activity (ROMA); .60 for Aimless motor activity (AMA);.81 for Distracted Attention (D); .72 for Looking away (LA); .96 for Verbalisation (V); . 59 for Vocalisations (Vo) and .91 for out-of seat (OS).

A series of bivariate Pearson's correlations ($\alpha = .05$, two-tailed) were used to explore the relationships between the adjusted frequencies per behavioural category and children's waiting performance on Day 1 (Table 6.3). In line with findings from the two previous studies, results indicated that children's overall levels of off-task and distraction activities as measured by the Overall Distraction Index were significantly negatively correlated with reward oriented behaviours even on the first day of training. Further to that and according to expectations, the mean Delay Time on Day 1 was strongly correlated with the duration of the free waiting trial and negatively correlated with children's self-reported delay aversion (although the latter relationship failed to reach statistical significance).

Table 6.3. Correlations between the adjusted frequencies per behavioural category and children's waiting performance on Day 1 of training.

		1	2	3	4	5
1	mDT Day 1					
2	Free Waiting Trial Day 1	.51*				
3	DASS-P-S Day 1	-.41	.12			
4	Adjusted RMA	.32	.07	-.36		
5	Adjusted ROMA	-.05	-.15	-.30	.01	
6	Overall Distraction Index	.12	.24	.24	.15	-.52*

Note: mDT: mean Delay Time; RMA: Restless motor activity; ROMA: reward-oriented motor activity.

6.4.6. Properties of the revised Delay Aversion Self-report Scale for pre-schoolers (DASS-P)

The internal consistency of the revised DASS-P scale was very satisfactory (Cronbach's alpha = .79). The scale's test retest reliability was also good. This was estimated as the correlation between the T1 and T2 DASS-P scores obtained for the waiting list group only (Pearson's correlation r = .65, significant at α = .01 level, two-tailed). For more information on inter-item correlations and item total statistics, please also see Table 6.4. A further correlational analysis was conducted to explore the scale's convergent validity with other delay tolerance measures. The results of this analysis indicated that children's self-reported delay aversion did not correlate with any other delay tolerance measures (Table 6.5).

Table 6.4. Inter-item correlations and item total statistics of the DASS-P scale

	Item 1	Item 2	Item 3	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 1. Do you like waiting for things	1.00	.67	.51	.66	.70
Item 2. How do you feel when you have to wait for things	.67	1.00	.55	.69	.65
Item 3. How hard do you have to try when you have to wait for things	.51	.55	1.00	.58	.79

Table 6.5. Correlations of DASS-P scores with all other pre-/post-training measures at T1

		1	2	3	4	5	6	7	8	9	10
1	DASS-P										
2	Teddies Score Mod	.14									
3	Teddies Comp.	.13	1.00**								
4	Bee Delay Mod	-.07	.50**	-.09							
5	Bee Mean Delay Comp.	-.03	-.05	-.05	1.00**						
6	SDQ Total	-.09	-.30	-.11	-.38*	-.31					
7	SDQ Hyperactivity	.10	-.20	-.25	-.35*	-.33	.78**				
8	QDQ Score	.15	-.16	-.13	-.28	-.28	.70**	.67**			
9	BRIEF-P ISCI	.05	-.36*	-.12	-.27	-.11	.75**	.73**	.71**		
10	BRIEF-P GEC	.02	-.30	-.08	-.32	-.19	.82**	.77**	.67**	.91**	

Note: Teddies Score Mod: Modified Teddies Score; Teddies Comp.: Teddies complete cases only; Bee Delay Mod.: Modified Delay Index (sum of total delay intervals); Bee Mean Delay Comp.: Mean delay per trial (complete cases only); DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

6.4.7. Training effects

A series of 2 (group) x 2 (time) repeated measures ANOVAs were conducted on all pre-/post-training outcome measures, with age and gender added as a covariates to the repeated measures design. Significant main effects of gender (as a between-subjects factor) emerged for three outcome measures, QDQ scores, $F(1,33) = 4.20$, $p = .05$, partial $\eta^2 = .11$, BRIEF-P total scores , $F(1,33) = 8.21$, $p = .007$, partial $\eta^2 = .20$ and BRIEF-P inhibitory self-control index , $F(1,33) = 8.67$, $p = .006$, partial $\eta^2 = .21$ and for this reason this covariate was retained in subsequent analyses. No significant main effects of age, or age by time interactions were obtained and as a result age was dropped as a covariate from the ANOVA models.

The descriptive statistics and results of these analyses are presented in Table 6.6. As can be seen in this table, neither the effect of group nor the interaction between group and time were statistically significant with regard to any outcome measure. There was however, a noticeable trend in the data for a group by time interaction effect in two variables: children's QDQ scores and their total delay time as measured by the Bee Task. In the case of the QDQ scores, the interaction effect was driven primarily by the larger gains noted in the performance of the intervention group. Notably, the group difference at T1 was nonsignificant but there was a group difference at T2. In the Bee task, the pattern of change over time was different for the two groups, with the intervention group's performance deteriorating from T1 to T2, while the control group's performance improving (note this trend in the data was not present when performance on the Bee Task was analysed for complete cases only). Again, the two groups did not differ in performance at T1, but they did at T2.

Table 6.6. Pre- and post-test performance (means and standard deviations) per group

Main effects analyses also included.

	Pretest (T1)		Posttest (T2)							
	Training Group	**Waiting List Group**	**Training Group**	**Waiting List Group**	**Group**			**Time**		
	M (SD) [N]	M (SD) [N]	M (SD) [N]	M (SD) [N]	F	p	$_p\eta^2$	F	p	$_p\eta^2$
Teddies Score Mod	13.68 (3.96) [19]	13.06 (3.98) [18]	13.26 (3.98) [19]	12.32 (4.12) [18]	F(1,34)=.28	.60	.01	F(1,34)=.34	.56	.01
Teddies Comp.	.49 (18) [17]	.50 (12) [13]	.46 (.20) [16]	.47 (.18) [11]	F(1,24)=.03	.86	.00	F(1,24)=.03	.88	.00
Bee Delay Mod	316.81 (194.03) [19]	296.59 (151.48) [18]	293.81 (179.54) [19]	363.26 (220.89) [18]	F(1,34)=.50	.49	.01	F(1,34)=.75	.39	.02
Bee Mean Delay Comp.	34.84 (17.91) [17]	38.77 (8.57) [12]	32.68 (17.67) [16]	47.42 (16.91) [12]	F(1,24)=2.27	.14	.09	F(1,24)=.06	.81	.00
DASS-P	.79 (1.47) [19]	1.39 (1.97) [18]	1.84 (2.06) [19]	1.94 (1.95) [18]	F(1,34)=.41	.53	.01	F(1,34)=.54	.47	.02
SDQ Total	12.76 (5.12) [19]	13.85 (4.82) [18]	11.86 (5.97) [19]	12.32 (4.15) [18]	F(1,34)=.11	.74	.00	F(1,34)=.87	.36	.02
SDQ Hyperactivity	6.84 (2.91) [19]	7.39 (2.52) [18]	5.81 (2.86) [19]	6.56 (2.33) [18]	F(1,34)=.30	.59	.01	F(1,34)=2.22	.15	.06
QDQ Score	27.89 (7.94) [19]	27.28 (7.49) [18]	20.90 (6.31) [19]	23.58 (5.42) [18]	F(1,34)=.05	.83	.00	F(1,34)=7.39	.01	.18
BRIEF-P GEC	103.58 (23.41) [19]	115.06 (28.08) [18]	102.39 (25.61) [19]	106.86 (21.26) [18]	F(1,34)=.51	.48	.01	F(1,34)=2.84	.10	.08
BRIEF-P ISCI	42.11 (10.13) [19]	47.39 (13.32) [18]	41.46 (11.11) [19]	44.58 (11.81) [18]	F(1,34)=.70	.41	.02	F(1,34)=2.57	.12	.07

Notes: Teddies Score Mod: Modified Teddies Score; Teddies Comp.: Teddies complete cases only; Bee Delay Mod.: Modified Delay Index (sum of total delay intervals); Bee Mean Delay Comp.: Mean delay per trial (complete cases only);DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

Table 6.6 (cont.)

Pre- and posttest performance (means and standard deviations) per group. Interaction effects analyses also included.

	Time by Group			Time by Gender		
	F	p	$_p\eta^2$	F	p	$_p\eta^2$
Teddies Score Mod.	F(1,1)=.30	.59	.01	F(1,1)=1.24	.27	.04
Teddies Comp.	F(1,1)=.04	.84	.00	F(1,1)=.11	.74	.00
Bee Delay Mod	F(1,1)=3.24	.08	.09	F(1,1)=1.43	.24	.04
Bee Mean Delay Comp.	F(1,1)=.18	.67	.01	F(1,1)=2.10	.16	.01
DASS-P	F(1,1)=.62	.44	.02	F(1,1)=.02	.88	.00
SDQ Total	F(1,1)=.21	.65	.01	F(1,1)=.09	.76	.00
SDQ Hyperactivity	F(1,1)=.15	.70	.00	F(1,1)=.36	.55	.01
QDQ Score	F(1,1)=3.68	.06	.10	F(1,1)=.66	.42	.02
BRIEF-P GEC	F(1,1)=1.53	.22	.04	F(1,1)=1.25	.27	.04
BRIEF-P ISCI	F(1,1)=.48	.49	.01	F(1,1)=1.45	.24	.04

Notes: Teddies Score Mod: Modified Teddies Score; Teddies Comp.: Teddies complete cases only; Bee Delay Mod.: Modified Delay Index (sum of total delay intervals); Bee Mean Delay Comp.: Mean delay per trial (complete cases only); DASS-P: Delay aversion self-report scale for pre-schoolers; SDQ: Strengths and Difficulties Questionnaire; BRIEF-P-GEC: Behavior Rating Inventory for Executive Function-Preschool Version- Global Executive Composite; BRIEF-P-ISCI: Behavior Rating Inventory for Executive Function-Preschool Version-Inhibitory Self-Control Index.

6.5. Discussion

The first aim of this randomised controlled trial was to evaluate the training efficiency associated with the adapted WG training task in a sample of children presenting with high levels of delay related difficulties. The results indicated that children with delay difficulties were able to train on the adapted Waiting Game and to sustain the delays presented to them successfully. Children's average delay time significantly improved between the first and the last day of the training, with a large effect size (r = .88) obtained for this comparison. Children were also able to sustain waiting on a voluntary basis, evidenced by the long waiting times recorded in the free waiting trials, possibly as a function of training as suggested by the significant association between training delay times and free waiting times. The amount of time children chose to wait voluntarily was significantly correlated and on average much longer than the mean delay time specified by their daily training schedule (the duration of the free waiting trial remained constant over time, however, contrary to the mean delay time which increased steadily from day to day). In addition, it is notable that most children were able to persevere with the training as evidenced by the very high trial completion rate and the very low dropout rate recorded in this trial.

With regard to children's observed behaviours, children were found to engage in significantly less reward-oriented behaviours over the course of the training, and in more off-task and self-distraction activities. These two types of attention focus (reward-oriented as opposed to distraction-oriented) were significantly negatively correlated even on the first day of training. In contrast, children's activity levels were not affected by the training and remained constant over time. Changes in behaviour over the course of the training suggest that children were able to spontaneously develop self-distraction to facilitate efficient waiting even without having been given explicit instructions on how to do that. As discussed in Chapter Five, these findings are still consistent with the basic tenet of the delay aversion model that inescapable delays may lead to children engaging in off task behaviours in an attempt to alter the subjective experience of delay (Sonuga-Barke, 2003). Seen from this perspective, the children of this trial did not become less delay averse during training (as evidenced by their increasing levels off task behaviours) but more waiting efficient (as evidenced by the fact they were able to sustain increasing delays successfully and see the training sessions to completion).

This pattern of results is also consistent with the classic literature on the effects of attention to rewards in the delay of gratification paradigm and fully replicates the findings of the second feasibility study. The negative association between children's mean delay times and their overall levels of distraction recorded on the first day of training replicates findings of an analysis of spontaneous attention deployment during a delay of gratification task (Peake, et al., 2002). Where our study goes further is exploring how this pattern of attention focus (towards or away from rewards) evolves over an extended period time.

It is also worthy of note that the frequencies of the observed behaviours obtained in the current trial were all elevated compared to the frequencies reported in the second feasibility study. This finding was in line with expectations given that the children recruited in the current trial were likely to find a structured waiting schedule more challenging than typical controls and to exhibit a wider range of behaviours in response to it. The fact the Observational checklist was sensitive enough to pick this trend up strengthens its psychometric profile. The results of the inter-rater reliability analysis add to it even further with the kappa statistic ranging from .59 to 1 in this sample, which was an improvement to the .49-1 range reported in the previous study.

The second aim of this small scale RCT was to investigate whether training efficiency transferred into training efficacy. That is, whether increases in waiting efficiency acquired during training generalised in other delay related behaviours, both proximal to the skills trained (e.g. as measured by similar choice delay tasks) and more distal (such as teacher-rated classroom behaviours). Unfortunately, the gains on these outcomes reported in the current trial were minimal and did not point to any differences between the two groups (no main effects of group or time by group interaction effects were accepted at the 0.05 alpha level). This pattern of results could be attributed to a number of factors: firstly, the intensity of the training may have been too easy for the children. Past studies have shown that training-related gains are easier to detect on difficult tasks (Diamond & Lee, 2011). This could potentially explain why children who trained for five days exhibited performance virtually the same as children who did not receive any training at all. Another possibility is that the two measurement points (T1 and T2) were so close to one another no detectable behavioural changes could have possibly occurred, especially in terms of global behaviours captured by the three behavioural rating scales.

Thirdly, it is also possible that improvements in performance were masked by fatigue effects. Indeed, some of the results obtained in the current trial may point to this explanation. Even though marginal improvements in executive functioning, hyperactivity and delay averse behaviours were reported by teachers (differences were not statistically significant, but teachers' scores did change in the right direction indicating small improvements), children's neuropsychological performance did not typically improve.

This result raises the recurring issue of the timing of the post-test and perhaps lends itself to yet another reference to the resource depletion hypothesis. The hypothesis put forward by Muraven and Baumeister (2000) views self-control as a limited resource which is amenable to practice. Crucially, the model allows for long term improvements in self-control "strength" after practice, if the "resource" is allowed to replenish following a period of rest (see also Persson, Welsh, Jonides, & Reuter-Lorenz, 2007 for the application of this hypothesis in the context of a cognitive training paradigm). The proposition that fatigue immediately follows practice only to be followed by increased "strength" at a later time could explain the seemingly puzzling finding of children making gains in training efficiency during training but then exhibiting (and self-reporting) delay fatigue at the time of the post training assessment. Future research should aim to manipulate the timing of the T2 assessment or add additional time measurement points.

Lastly, the possibility should always be entertained that behavioural changes incurred in the context of such a structured learning/training paradigm may not readily generalise to other related constructs, particularly after such a brief period of training. The majority of play-based interventions for preschool ADHD (see for example Halperin et al., 2012; or Healey & Halperin, 2014) have durations counted in weeks (5-

6 weeks seems to be a standard intervention length). Delay shaping and fading interventions have always been lengthy as well (Schweitzer & Sulzer-Azaroff, 1988) and have scarcely looked into transfer effects of training. It is also entirely possible that adopting a more localised experimental approach and focusing on one training task for a limit period of time is not a paradigm suited to reward processes. Changes in motivational preferences may take a longer time to establish than changes in higher cognitive processes and may not be detectable within the timescale set out by the process-specific training paradigm.

Given all the above, assessing the feasibility of the WG training task as an intervention for children with ADHD becomes both a straightforward and at the same time extremely complicated task to address. The straightforward answer to this question is that inasmuch practice on the WG training task did not yield any generalizable training effects the wider dissemination and full scale evaluation of the programme in its current form cannot be recommended. What is more complicated is the broader assessment of the potential value of delay training interventions for pre-schoolers with ADHD on the basis of the current findings. As a means of approaching this wider question, one may begin by listing the successes and failures of the training paradigm implemented in this trial.

In terms of its success, the WG training programme demonstrated that children can learn to sustain longer and longer delays in order to maximise their rewards while spontaneously learning where to focus their attention in order to achieve this goal. Children were also highly motivated to take part in the training study and also to persevere with it until the end, so in terms of acceptability and completion rate the adapted version of the Waiting Game worked well.

The WG paradigm was not able to demonstrate how the gains in waiting efficiency recorded during training may generalise to other related constructs (e.g. reward preference in slightly different contexts) and also affect children's behaviour (or functional impairment in the case of children with ADHD) more broadly. Further works is needed to address issues specific to the nature of delay training (such as the timing of the post-training assessments) as well as the overarching issue of reward and delay fatigue as potential moderators to treatment response.

6.5.1. Limitations

The current trial has important limitations: firstly, as was the case with the previous feasibility studies, the fact this was a small-sized RCT meant it lacked the statistical power necessary to provide definitive answers to some of the questions it asked and, further to that, to explore many other issues of interest, such as individual differences to treatment response and potential mediators and moderators of it. The use of a waiting list control group and resulting lack of blindness to group allocation may have resulted in expectation bias or inflated ratings by teachers (although, given the pattern of findings obtained in this trial, this factor did not have a detectable effect in this case). In addition, another limitation is the lack of parent data. Parents would be more likely to be blinded to the potential effects of a school-based intervention and as a result a more unbiased source of information and should be included as informants in future delay training studies.

Chapter 7: General Discussion

7.1. Aim of this chapter

The aim of the chapter is to provide an overview of the empirical findings presented in this thesis and to evaluate the effectiveness of the Waiting Game intervention in terms of its potential theoretical and practical implications for the management of delay related difficulties in clinical and educational contexts.

7.2. Summary of results

This thesis has presented a series of studies that have evaluated the feasibility of delay training as an intervention for typically developing and young impulsive children at risk for ADHD. It has also developed a number of novel methodological approaches to measuring delay related behaviour in preschool aged children. This was an essential step in the process of assessing the potential of the delay training paradigm to effect behaviour change.

7.2.1. Delay training findings

The main findings derived from the delay training studies presented in Chapters Five and Six can be summarised in schematic form as follows:

Children learned to wait efficiently during training. In other words, children were able to wait as long as they needed in order to maximise the rewards offered by the training task. This was demonstrated by comparing the average waiting time children achieved on their first day of training to that achieved on their last day. Indeed, results indicated that children's average delay time significantly improved between the first and the last day of the training, with large effect sizes ($r > .8$) obtained for this comparison in a typical sample (Study Two) and in a sample of children with elevated levels of delay difficulties (Study Three). A similar trend was also present in the first study as well, but the effect was compounded by the low completion rate observed in that study.

Children were able to persevere with the training task, when rewarded specifically for persevering (i.e. for bringing trials to completion). In the first training study, typically developing children found it difficult to persevere until the end of the training sessions, particularly as the delay intervals grew longer during the final days of the training. To address this issue a new reward contingency was put in place offering children an additional valued reward if they managed to wait until the end of the delay periods. This manipulation was very effective and led to high rates of completion during the training in subsequent studies. It is interesting to note, however, that this manipulation was not applied in the case of the delay tasks administered as part of the pre- and post-training assessments where low completion rates were again reported in Study 3 (which used a sample of impulsive children). The issue of seeing delay tasks to completion is an important one because anecdotal evidence suggests young children may use early exit from delay tasks as a delay avoidance strategy after failing to make full use of the delay escape mechanisms inherent in these tasks, i.e. favouring the smaller-sooner options.

Children spontaneously adapted their waiting behaviour during training. Data obtained using a novel and reliable observational system indicated that as children learned to sustain longer waiting times, they also learned to switch their focus away from the rewards and towards other distractions without explicit instruction. This finding was consistent across all the three training studies and was further corroborated by correlational analyses which revealed that better delay times on the first day of training were negatively correlated with higher levels of reward oriented activity. This latter finding was also very robust and was obtained across all training studies.

Gains in waiting efficiency did not transfer into generalised improvements in delay tolerance/waiting outside the training setting. This series of studies investigated whether increases in waiting efficiency reported during training generalised to other constructs, both proximal to the skills trained (e.g. motivational preferences as measured by choice delay tasks) and distal (such as delay related classroom behaviours). Unfortunately, the gains on these outcomes were minimal and did not point to any differences between the training and comparison groups across studies.

7.2.2. Methodological innovation

This thesis embarked on the development and appraisal of a number of innovative methods for measuring delay sensitivity in very young children. This is still a work in progress as some of these measures will benefit from further refinement, but several important advances have already been made. These can be summarised as follows:

A novel observational system was developed as a complement to the delay training task. Its aim was to capture the behaviours children engaged in while waiting. This measure underwent a protracted development process and the behavioural categories it focused on were refined over the course of several iterations. The end result of this process was an observational protocol with excellent psychometric properties (satisfactory inter-rater reliability and good external validity) which highlighted aspects of delay training performance that would not have been easy to capture in any other way. In fact, without the use of observational methods, it would have been very difficult to make any real contribution to theory or provide any arguments as to why children performed in the way that they did. It was the use of observational methods that allowed us to formulate an account not only of the experimental results but also of why this pattern of results may have occurred.

An age sensitive neuropsychological delay task was designed, piloted and implemented. The Bee Delay task is a novel neuropsychological measure of delay tolerance designed specifically for young children. As discussed extensively in Chapter 4, the Bee Delay task was shown to be a more sensitive and more flexible index of delay tolerance compared to existing choice delay tasks, as it calculates a continuous index of the total amount of time young children are willing to wait for delayed rewards. It also uses an intuitive set up to make delay choice, i.e. the trade-off between reward magnitude and delay, easier for children to understand. The findings of an initial pilot study provided preliminary evidence for satisfactory reliability, validity and age sensitivity of the task. Furthermore, the Bee Delay task was the only neuropsychological task which captured differential performance between the two groups in Study Three, providing further evidence of its increased sensitivity.

Teacher ratings of delay aversion. This thesis also initiated the process of validating an existing delay aversion questionnaire as an instrument that can be used by teachers and caregivers to evaluate children's delay related behaviour along two dimensions, delay aversion and delay discounting. A small validation study used a pooled dataset from all the studies completed as part of this thesis and indicated that the teachers' version of the QDQ scale was reliable, had internal consistency and was highly sensitive to clinical problems, specifically ADHD and CD. However, further research is needed to corroborate these findings as this analysis was done on a relatively small sample of young children and used teacher ratings only.

A self-report delay aversion scale for young children. This was a three-item scale asked children three questions: whether they liked waiting, how they felt when they had to wait and whether they found waiting easy or difficult. To answer these questions children were asked to point to pictures visualising different response options. The revised version of the scale (Study Three) had good internal consistency and test-retest reliability. However, it showed poor external validity, as it failed to correlate with any other delay tolerance measure. It is hard to know at this stage whether this is due to the design of the scale or a reflection of the fact children's perception of their waiting performance differs from their actual waiting performance to a considerable degree. The issue of biased self-evaluations in children with ADHD is a known one (Owens, Goldfine, Evangelista, Hoza, & Kaiser, 2007) and worth exploring further in the context of sensitivity to delay, particularly in view of the emotive element of delay aversion which is clearly of a subjective and introspective nature. A similar approach was employed recently with a sample of school-aged children with ADHD, who were asked to rate their difficulty waiting on a visual analogue scale (Scheres, Tontsch, & Thoeny, 2013). Results were unexpected in that the association between subjective and objective measures of difficulty waiting were found to be higher in the ADHD group than in the control group. The authors of the study suggested this could be an indication that typically developing children also felt waiting was difficult for them, but were able nonetheless able to override these feelings and wait. This is a further demonstration of why a mixed methods approach to these questions is not only informative and but essential to hypothesis testing and further theory development in this field.

7.3. Why did gains in waiting efficiency not generalise into improvements in delay tolerance?

Null findings in experiments are often as important in driving scientific progress as positive findings. This is because they help reject parts of current theory that are untenable in view of the new results and open the way to generating new theory which is a better fit to the empirical data. This is also the reason why it is crucial to publish null results in the literature: they are an essential step in the process of new theory formation and a well-described part of the scientific process (Kuhn, 1962). Where the challenge lies for the individual researcher is in being able to determine when a null effect has been obtained.

As has already been pointed out in the discussion sections of the respective studies, there are a number of contextual factors that may have affected the impact of the Waiting Game intervention as assessed in a repeated measures design. In summary form, these are: firstly, some aspects of the intervention may need further adjustment, particularly the intensity and/or the duration of the training which may have been too easy or too brief for it to produce any meaningful treatment effects. Secondly, the timing and the spacing of the T1/T2 measurement points may be crucial. It is possible that in the studies described in this thesis the measurement points were so close to one another (two weeks apart for the neuropsychological measures and four weeks for the behavioural ratings) no detectable behavioural changes could have possibly occurred in that time. Thirdly, it is also possible that improvements in performance were masked by fatigue effects. Indeed, some of the trends in the data may point to this explanation, with children in the training group of Study Three reporting more delay aversion and performing worse in the Bee Delay task than children in the waiting list condition. This is a very important issue for future studies to resolve as fatigue effects may impact greatly on children's T2 delay performance. If practice on delay tasks causes fatigue, it is important to establish the nature of these effects (do they affect both subjective and objective measures of delay tolerance to the same degree, for example?) and their persistence across time. The potential theoretical implications of these effects are also worthy of mention, as they seem to impinge on the self-regulatory resource depletion model (Muraven & Baumeister, 2000) and could provide a new paradigm for testing the predictions of this model using multiple applications of the same task, instead of employing two or three different tasks (as in Converse & DeShon, 2009 for example). Fourthly, the question needs to be asked whether the picture would have been any different if different outcome measures had been used. Every effort was made to select primary measures appropriate for this age group, but given the novelty of the research design and the many measurement issues surrounding the neuropsychological testing of very young children the full range of the training effects may have remained undetected.

All the above contextual factors notwithstanding, the possibility should also be entertained that this trial did in fact produce a “real” null intervention effect and that behavioural changes incurred in the context of a structured delay training paradigm did not readily generalise to other related constructs. This interpretation of the findings has several theoretical implications, which will be explored in detail in the following section.

7.4. Are these results consistent with existing models of impulsive choice?

Assuming that after practicing on a delay training task children did learn to wait more efficiently, but gains in waiting efficiency did not generalise into further improvements in delay tolerance, how do these findings fit with current models of impulsive choice?

Changes in the observed behaviours of children over the course of the training are consistent with the classic literature on the effects of attention to rewards in a delay of gratification context (Mischel, et al., 1989; Mischel, et al., 1972). The negative association between children’s mean delay times and their overall levels of distraction recorded on the first day of training replicates findings of an analysis of spontaneous attention deployment during a delay of gratification task (Peake et al., 2002) and provides further support to the idea that an attentional focus on rewards undermines children's ability to wait successfully. Where our studies go further is in exploring how this pattern of attention focus (towards or away from rewards) evolves over an extended period time. Mischel and colleagues have also demonstrated that compared to older children four-year-olds are very ineffective in directing their attention away from rewards, preferring instead to look at the rewards and to think about them (Mischel, et al., 1989). In contrast, our studies have suggested that given enough time and the opportunity to practice, even children as young as four may be able to develop self-distraction strategies that facilitate effective waiting. Importantly, our children were able to generate these strategies spontaneously and without having been given any instructions or prompts to that effect. This is an important finding for this thesis as it suggests that children are able to develop self-regulatory patterns of coping with waiting situations spontaneously if they are given the opportunity to practice waiting. Interestingly, the coping strategies children came up with were very similar to those older children often self-generate while waiting during a DG task.

Changes in children's observed behaviour over the course of the training is also consistent with the concept of delay aversion, as originally developed by Sonuga-Barke (Sonuga-Barke, 2003; Sonuga-Barke, et al., 1994; Sonuga-Barke, Taylor, Sembi, et al., 1992). A basic tenet of this model is that delay aversion manifests primarily in conditions where delay is unavoidable. In these situations, it is hypothesized that children's difficulties with waiting lead them to direct their attention to aspects of the environment that can help them speed up the passage of time. These behaviours can range from fidgeting to going off task and engaging in other self-distraction activities. In the context of the DA model all these behaviours are construed as compensatory strategies used to help children escape the subjective experience of delay. In this light, children's tendencies to engage in increasingly higher levels of self-distraction as the training progressed in the current studies could be seen as an attempt on their part to transform the experience of delay through various coping mechanisms (e.g. some children may have engaged in aimless repetitive movements while others may have resorted to talking to the experimenter more).

It is important to note that these behaviours were elicited in a situation where delay was not strictly speaking inescapable (as children could have cut the delay trials short by forgoing some of the rewards), which is not entirely compatible with the predictions of the DA model. But it appears the majority of the children who underwent the training were highly motivated to get to the end of the delay periods and in doing so they generated strategies that helped them wait successfully. In this sense, therefore, our results take the DA model a step further by suggesting that even in conditions of self-imposed delay (where delay is escapable, but children prefer to stay and wait) children will still try to find ways to modify how delay feels. To sum up so far therefore, children’s increasing use of self-distraction strategies during the course of the training is consistent with both the DG and DA models although to accommodate these findings the DA model would need to make allowances for situations where delay is escapable but self-imposed (i.e. there is a strong desire to wait for a valued outcome).

The finding that gains in waiting efficiency did not generalise to improvements in delay tolerance is harder to interpret with reference to current models of impulsive choice. Sonuga-Barke’s exposure model (2004) predicted that exposure to delays of increasing length should reduce delay aversion. The null intervention effects obtained in this thesis do not support this prediction. There are two possible explanations as to why this might be so. The first would be to concede that the experience of delay cannot be altered merely by exposure training. The second would be to argue that training by delay exposure was not implemented successfully in this training paradigm because children could still escape delay as they trained.

The main objective of the WG was the creation of positive delay related experiences through practice on repeated instances of achievable, rewarded delay trials. This was thought of as a desensitization process through which repeated exposure to brief instances of delay would lead to diminished (negative) emotional responsiveness to delay. It was also thought of as a way of strengthening the association between a response and its delayed (positive) reward outcome, which is considered to be weak in ADHD. During training, and in line with findings from prior research, children were allowed to apply self-distraction strategies as these have been repeatedly shown to enhance self-regulation. The rationale was that the main objective of the training paradigm was to help children achieve rewarded delay in any way that they could and self-distraction is one of the ways children use to achieve efficient waiting. However, and in view of the null transfer effects obtained in this thesis, the manipulation of allowing children to self-distract while waiting may have led to unanticipated results.

Allowing children to disengage their attention from their rewards may have enabled them to learn how to wait efficiently but it may have also prevented them from being exposed to the most aversive aspects of waiting, which in turn may have undermined the very idea behind the delay exposure model. This observation leads to an interesting conundrum as another basic tenet of the model was that the delay experiences offered to the children must be successful (i.e. rewarded) and hence easy, pleasant and achievable at least at the beginning stages of the training. The question therefore becomes if one could ever arrange a situation where delay exposure can work after all these considerations are taken on board, so that it leads not only in better waiting efficiency but also in more generalised improvements in delay tolerance. A few speculative ideas on this theme will be explored in the next section.

7.5. Is delay exposure valuable as a therapeutic approach?

For delay exposure to work as a training approach, it might be necessary to expose children to waiting situations that are particularly aversive (e.g. experimental conditions where they will be asked to wait without being able to apply any self-regulation strategies). Interestingly, a novel method of assessing self-control in preschool children has recently been proposed which makes very similar demands. In the new Watch-and-Wait Task (WWT), preschool aged children are offered a LL reward (a second toy) only if they are able to wait without looking away or talking; instead they are instructed to watch an hourglass run out as they wait. Children receive two warnings if they break the looking-or-talking rule and the task is stopped if they break the rule for a third time resulting in them receiving the SS reward option (one toy). Performance on the WWT task has been shown to have reasonable test-retest reliability and good association with later academic achievement and behavioural problems (Neubauer, Gawrilow, & Hasselhorn, 2012). Nevertheless it is hard to imagine how waiting under such harsh conditions could be readily converted in a training/learning paradigm. As it has already been emphasized in this thesis enabling young children to persevere with delay tasks is no easy task. Delay tasks featuring particularly aversive waiting conditions will no doubt require dispensing very large rewards (possibly monetary or edible ones) to very young children which will automatically render them impractical as an intervention and very likely unpalatable to both educators and parents.

This should not mean, however, that delay exposure does not have any value as a therapeutic mechanism. As it has been discussed extensively in the introductory chapters of this thesis impulsive choice is an excellent target for early intervention and more effort should be put into developing tasks that incorporate delay exposure in order to enhance early delay ability. Such tasks could manipulate delay difficulty by not only increasing the length of delay intervals, but also by varying the type of waiting conditions. For example, it would appear that tasks that allow participants to set their own reward magnitude targets, such as the Bee Delay task, are a good way to introduce children to the concept of practicing waiting and to strategies that may help them achieve their waiting targets more efficiently. Tasks should then become progressively more challenging by exposing children to longer waiting times and less flexible delay situations (e.g. of non-escapable delay or where self-distraction is inhibited) while at the same time helping children select the most appropriate strategies for each challenge (e.g. cognitive reframing could be taught as the best strategy to use when faced with tasks that restrict other types of self-distraction). The use of an extended battery of delay tasks (e.g. resistance to temptation tasks, self-adjusting tasks, choice delay tasks, etc.) in conjunction with an array of self-regulation strategies should help children apply their learning across contexts and thus potentially lead to more generalised and transferable training effects. It should also serve to make such activities more palatable to children over longer periods of time. Changes in motivational preferences may take a longer time to establish and children's continuous engagement with any such programme of activities could be one of the hardest obstacles to overcome.

The issue of whether the behaviours children engaged in while training on the WG task were adaptive or maladaptive also needs to be considered. Children made increasing use of self-distraction to help them maximise their rewards. Clearly, that was an adaptive strategy for them and to the training setting. The question remains however, whether the strategies children generated and found successful in the context of delay training would have been maladaptive if they had been used in other settings. For example, it is easy to see how excessive fidgeting or going off task helped children sustain long waiting periods during the training but such behaviours may not have been particularly adaptive had they been used in the classroom. To assure delay ability is enhanced in ways that can be used adaptively in the classroom it appears that self-regulation strategies may be a particularly useful tool, which could be used in conjunction with delay exposure tasks. For example, self-regulation strategies designed to address particular waiting situations children may find difficult in class could make use of adaptive coping strategies children spontaneously generate during delay training. The participation of parents and teachers in the identification of these strategies should also help with the successful implementation of any such intervention programme.

7.6. Future Directions

On the basis of the empirical findings obtained in this thesis and their implications, a number of suggestions can be made for future research:

The measurement of difficulty waiting needs to be extended to include mixed and multimodal research methods: these can include observational measures of behaviours while waiting, measures of the subjective feelings of the experience of waiting and of one's efficiency at waiting (e.g. whether one's own initial waiting intentions have been realised by the end of the delay periods) as well as neuroimaging methods that could help clarify these aspects of waiting behaviour further.

More work is needed before a hierarchy of waiting situations in order of difficulty can be sketched out. This work is of paramount significance as working at the correct level of difficulty is a prerequisite for any training programme. It is also crucial in view of the numerous contextual and individual state factors motivational preferences can be influenced by.

The issue of fatigue but also of replenishment effects on delay performance also needs further disambiguation. If practice on delay tasks causes fatigue and hinders delay performance in the short term, it is important to establish the nature of these effects, e.g. does practice eventually enhance performance after a period of "recess"?) and their persistence across time. A related unexplored area is that of the time scale of training. It is possible that to strengthen self-control skills many hours of practice over extended periods of time are required, making this type of training difficult to implement as a psychological intervention and more appropriate to embed into academic curricula. Future training studies employing carefully timed follow up measurements should be able to address these issues.

Future studies should also use larger samples of children that will allow the investigation of individual differences to treatment response and potential mediators and moderators of intervention effects.

Finally, the issue of the most appropriate mode of delivery for these interventions also needs some attention, particularly with regards to interventions targeting young children. Because of the extensive tailoring delay exposure interventions need, their most common mode of delivery is through individualised programmes. But findings indicate that interventions that are more ecologically valid have been more effective in this age group. Further to that, the costs and resources associated with delivering individualised training programmes hinder their implementation as preventive early interventions. Further research should focus on the challenges associated with delivering these programmes in group settings or as part of Early Years curricula, including sources of individual differences, age-dependent scaling of the programmes and group and peer effects on the expression of difficulties with waiting.

Concluding remark. Interest in ways to teach children to sustain delay in exchange for longer term goals spans different fields of psychological research and many decades. Still, the question of whether this skill can be taught is a largely unanswered one. And despite large parts of parenting and Early Years education efforts coming together in trying to enhance it, it staggering to think how little effort is currently expended by researchers attempting to find out more about the mechanics of teaching such an important skill to young children. Clearly more research is needed to shed light to the multifaceted nature of this question.

Appendix A.1 The Bee Delay Task: Description and user guide

The Bee Delay Task: Description and user guide

1.1 Description of the task

The Bee Delay ask is an adjustable delay task. It is designed to measure children's delay tolerance by calculating the exact amount of time children are willing to wait for delayed rewards.

The task instructs children that in this game a bee is going to help them win points. At the beginning of each trial, children are asked to choose the number of flowers they want to put on the screen for that bee to stop at. For each flower the bee stops at they will win one point. But it is explained to them the more flowers a bee stops at, the more tired it gets and the longer it takes to fly to the next flower. Children are also told that they have a stop button they can use if they think one of their bees has become too slow. If they press it, they can move on to a different bee and start over, but they will only get the points they have won before aborting the trial.

Each trial includes a selection screen and the actual play screen:

- *Selection screen.* In this screen, children are shown a landscape scene with a fixed number of positions (flower stems) where they can put flowers on, by clicking on them. They are told they can pick up to seven flowers for their bee to stop at (please see paragraph 1.5. for instructions)

- *Play screen*. After the experimenter has pressed the Enter key the trial starts and a bee appears on the left hand side of the screen. When the bee flies over a flower, a buzzing sound is heard, the flower disappears and one point is reserved at the bottom of the screen where a faded (grey) flower becomes visible. All reserved (grey) points remain on the screen for the duration of the trial. When the trial is finished the reserved (grey) flowers turn bright yellow and the corresponding rewards are dispensed. If the trial is aborted before all stops are achieved (by pressing the space key), then only grey points are won (it is up to the experimenter to dispense those or not).

1.2 Delay interval timings

The lengths of the between-stops delay intervals are determined by two parameters:
1. Total delay interval (TDI) of the longest possible trial (when the maximum number of stops is selected)
2. Rate (a) at which each of the delay intervals gets progressively longer as the trial progresses

$DI_{1...7}$: between-stops delay intervals (the time a bee takes to move from one flower to the next)
$TDI_{1...7}$: total delay interval of each trial (total flying time of the bee per trial)
DI_c : the interval between the last stop (last flower achieved) and the termination of the trial. This constant will be set to a duration of 1 sec.

$DI_1 + DI_2 + ... + DI_7 + DI_c = TDI_7$

$DI_2 = aDI_1$
$DI_3 = aDI_2 = a^2DI_1$
$DI_4 = aDI_3 = a^3DI_1$
...
$DI_7 = aDI_6 = a^6DI_1$

1.3. Outcome measures and stats

On each trial, the number of stops per trial (i.e. the number of flowers the child selects) and the total DI are recorded. If the child uses the stop button to terminate the trial, the number of achieved stops is recorded and the duration of the achieved TDI as well. Please also see table below.

	Number of stops selected by child (Flowers chosen)	Projected total delay interval (TDI) (on the basis of Flowers chosen)	Actual number of stops achieved (Flowers Achieved)	Actual TDI(trial time until stop button pressed)
Trial 1				
Trial 2				
…				
Trial 10				

Statistics: sums and means (across the ten trials) are recorded in the Bee Task output file for all four columns. To produce outcome variables comparable to the most frequently used delay task outcome, namely the ratio of LL responses over the total number of delay trials, the two main indices of the Bee task (total number of flowers achieved, and the Actual TDI sum) can also calculated as ratios (over the total number of achievable flowers = 70 and over the maximum possible TDI = 900sec). These ratios are not computed by the output file itself, but can be easily calculated at the data processing stage.

Two further indices are calculated by the output file, as a measure of the discrepancy between selected and achieved delay targets. One discrepancy index focuses on the number of flowers chosen and the other on the TDI. The former is calculated by dividing the actual number of stops achieved by the number of stops selected, and the latter by dividing the actual TDI by the projected TDI.

E.g. the discrepancy index for number of stops calculated for a game where a child has selected all seven flowers across all ten trials (hence 70 flowers in total) but has only managed to win 30 points would be 30/70 = 0.43

1.4. *BeeDelay Task Software*

The BeeDelay task software can be downloaded by double-clicking on the BeeDelay folder and then following the installation instructions. Once installed, click the BeeDelay icon on your start menu. This will start the task.

The first screen is the set up screen, where the following parameters can be set: number of practice and experimental trials, duration of the longest delay interval and rate of delay increase from one trial to the next. We recommend the following settings for children aged 4-5 yrs.

- Number of practice trials: 3 (also see paragraph 1.5 for instructions)
- Number of trials: 10
- End delay (please set value at 1.00)
- TDI7 (total delay interval of the trial where the maximum number of stops is selected): 90 sec
- α (rate of delay increase): 1.25

Once you have selected these parameters, you can click the start button to begin the task. This action will cause a pop up window to appear asking you to name the data file where all the data from the current session will be saved. Once you have done that and saved the file the task will move to the first selection screen.

Flowers can be selected by clicking on the empty positions (flower stems). The experimenter initiates the trial by pressing the Enter key. Children can abort the trial at any time by pressing the Space key. We recommend a sticker like the one shown below is placed on the keyboard to remind children of the position of the stop key. You can terminate the task at any moment by pressing Alt + F4.

1.5. *Task instructions*

"In this game, a bee is going to help you win points.

First you have to choose how many flowers you want your bee to stop at. For every flower the bee stops at, you will win one point. Let's see how this works (the researcher selects one flower and initiates the first practice trial).

After the trial is finished*: If you want to win more points, you can pick more flowers. But the more flowers a bee stops at the more tired it gets and the longer it takes to fly* *to the next flower. Let's see how long it will take if we choose all 7 flowers* (the researcher selects seven flowers and initiates the second practice trial).

After the second trial is finished, the researcher selects 7 flowers and initiates a third practice trial: *Now if you think your bee has become too slow, you can press this button to make the bee stop. If you make the bee stop, you will still get the points you have won until then. Let's see what happens when I press the button* (the researcher aborts the trial after 4 flowers have been won). *See, the bee stopped and you have won four flowers.*

Would you like to try yourself now? How many flowers would you like to start with?

Appendix A.2 Strengths and Difficulties Questionnaire

Strengths and Difficulties Questionnaire

T[4-16]

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 or this school year.

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?

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

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

No	Yes- minor difficulties	Yes- definite difficulties	Yes- severe difficulties
☐	☐	☐	☐

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

- How long have these difficulties been present?

Less than a month	1-5 months	6-12 months	Over a year
☐	☐	☐	☐

- Do the difficulties upset or distress the child?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

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

	Not at all	Only a little	Quite a lot	A great deal
PEER RELATIONSHIPS	☐	☐	☐	☐
CLASSROOM LEARNING	☐	☐	☐	☐

- Do the difficulties put a burden on you or the class as a whole?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

Signature .... Date ....

Class Teacher/Form Tutor/Head of Year/Other (please specify:)

Thank you very much for your help

Appendix A.3 Strengths and Difficulties Questionnaire (Follow-up)

Strengths and Difficulties Questionnaire

T4-16
FOLLOW-UP

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 month**.

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?

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

Since coming to the clinic, are the child's problems:

Much worse	A bit worse	About the same	A bit better	Much better
☐	☐	☐	☐	☐

Has coming to the clinic been helpful in other ways, e.g. providing information or making the problems more bearable?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

Over the last month, has the child had difficulties in one or more of the following areas: emotions, concentration, behaviour or being able to get on with other people?

No	Yes- minor difficulties	Yes- definite difficulties	Yes- severe difficulties
☐	☐	☐	☐

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

- Do the difficulties upset or distress the child?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

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

	Not at all	Only a little	Quite a lot	A great deal
PEER RELATIONSHIPS	☐	☐	☐	☐
CLASSROOM LEARNING	☐	☐	☐	☐

- Do the difficulties put a burden on you or the class as a whole?

Not at all	Only a little	Quite a lot	A great deal
☐	☐	☐	☐

Signature .... Date ....

Class Teacher/Form Tutor/Head of Year/Other (please specify):

Thank you very much for your help

Appendix A.4 Quick Delay Questionnaire

Quick Delay Questionnaire

Instructions: Please circle a number below which best describes your behaviour over the past 6 months, from 1 = 'Very like me' to 5 'Not at all like me'

		Very like me ←→ Not at all like me				
		1	**2**	**3**	**4**	**5**
1	Even if I have to wait a long time for something I won't give up if it is important to me	1	2	3	4	5
2	I am usually calm when I have to wait in queues	1	2	3	4	5
3	I will often chooses a task because it is beneficial in the long term even if it doesn't offer immediate reward	1	2	3	4	5
4	I feel relaxed when waiting for things	1	2	3	4	5
5	I often give up on things that I cannot have immediately	1	2	3	4	5
6	I hate waiting for things	1	2	3	4	5
7	I try to avoid tasks that will only benefit me in the long term and don't have any immediate benefits	1	2	3	4	5
8	I feel frustrated when I have to wait for someone else to be ready before I can do something	1	2	3	4	5
9	Having to wait for things makes me feel stressed and tense	1	2	3	4	5
10	The future is not important to me, I only consider the immediate consequences of my actions	1	2	3	4	5

Citation: Clare, Sylvia, Helps, Suzannah and Sonuga-Barke, Edmund J.S. (2010) The quick delay questionnaire: a measure of delay aversion and discounting in adults. *ADHD Attention Deficit and Hyperactivity Disorders*, 2(1), 43-48

Appendix A.5 Quick Delay Questionnaire – Parent and Teacher Form

Quick Delay Questionnaire-Parent and Teacher Form

Child's name/ID: ______________________ Age: ________ School Year: ______

Birth Date: ______/______/______ Today's Date: ______/______/______ Gender: M F
Day Month Year Day Month Year (please circle one)

Your name: ___________________________

Instructions:
The items below describe some types of behaviour children may exhibit. Please rate the following items from
1: 'Not at all like him or her' to **5: 'Very much like him or her'**.
Please try to rate all items.

		not at all like him/her				very much like him/ her
1	Will not give up, even if he or she has to wait a long time for something important to them	1	2	3	4	5
2	Usually calm when having to wait in queues	1	2	3	4	5
3	Will persevere with tasks even if they do not offer immediate rewards	1	2	3	4	5
4	Seems relaxed when waiting for things	1	2	3	4	5
5	Often gives up on things he or she can't have immediately	1	2	3	4	5
6	Hates waiting for things	1	2	3	4	5
7	Tries to avoid tasks that don't have any immediate benefits	1	2	3	4	5
8	Seems frustrated when asked to wait before he or she can do something	1	2	3	4	5
9	Fidgety and restless when having to wait for things	1	2	3	4	5
10	Does not consider or understand the future consequences of his or her actions	1	2	3	4	5

Appendix A.6 Quick Delay Questionnaire – Instructions Sheet

[The following instructions are the same for the QDQ Self-Report Form and the QDQ Parent and Teacher Form]

The Quick Delay Questionnaire (QDQ) is a 10-item scale. It includes two 5-item scales: (i) the Delay Aversion (DA) and (ii) the Delay Discounting (DD) scales.

Items 2, 4, 6, 8, 9 are included in the DA scale and the remaining items (1, 3, 5, 7, 10) in the DD scale.

A higher QDQ score indicates more difficulties in responding to delayed rewards and an exaggerated degree of delay discounting.

For the current version of the scales, where items are rated as **1: 'Not at all like me/him or her' to 5: 'Very much like me/him or her'**, QDQ scores should be obtained after items 1-4 are reverse coded.

Please note that in previous versions of the scale, the rating scale was reversed and informants were asked to rate items as 5: 'Not at all like me/him or her' to 1: 'Very much like me/ him or her'. For these versions of the QDQ, items 5-10 should be reverse coded before scoring.

We also suggest that before scores are obtained the scale is shifted to range from 0-4 (instead of 1-5). If this is done, the DA and DD scales should each range from 0-20 (instead of 10-30) and the aggregate QDQ scores should range from 0-40 (instead of 10-50).

Citation: Clare, Sylvia, Helps, Suzannah and Sonuga-Barke, Edmund J.S. (2010) The quick delay questionnaire: a measure of delay aversion and discounting in adults. ADHD Attention Deficit and Hyperactivity Disorders, 2(1), 43-48

Appendix A.7 Observational Checklist v2

The underlined sections represent updates from the previous version.

- *Restless motor activity (RMA)*. The child displays gross motor movements such as tapping a foot, rocking a chair back and forth, fidgeting, etc.
- *Aimless motor activity (AMA)*. The child displays aimless, waiting specific movements while waiting (touching hair, playing with clothes or jewellery, etc.). Sometimes these movements may accompany gross body movements. If this happens, both codes (RMA and AMA) are recorded. Ordinary movements such as changing body or hand positions or pointing and gesturing while speaking are not recorded as AMA, unless they are very repetitive or done in an exaggerated manner.
- *Reward-oriented motor activity (ROMA)*. Child leans towards and touches or lifts up the covers or in fact takes the reward stickers when inappropriate. Sometimes it is very hard to differentiate RMA/AMA behaviours from ROMA because the child displays, for example, aimless movement while at the same time leans towards the sticker covers. In such instances, all relevant codes are used. Note that if the ROMA code is used, then the trial is always scored as a partially inhibited (or as not inhibited, if the child actually retrieves the sticker). <u>If the child touches the sticker covers, then the trial is terminated and the child is refused the second sticker. If other ROMA-type behaviours are exhibited, then the ROMA code is recorded (and the trial is scored as 1) but the trial is allowed to time out and the child receives the second sticker. Note that if the child hovers over or even touches the sticker cover during the last 2-3 seconds of a delay interval, in anticipation of the end of the trial, then the ROMA code should be used but the trial should not be terminated and the second sticker should not be withdrawn.</u>
- <u>*Distracted Attention (D)*. Child plays with the stickers she has won so far, the dice, the stopwatch or other objects while waiting. This code aims to capture attempts by the child to distract herself while waiting using the experimental materials. But it is possible that other objects, e.g. a toy that the child has brought along, can be used as distractors as well. But note the difference between fiddling with a bracelet (which should be coded as AMA) and taking a bracelet off to show to the experimenter and use as a talking point (which should be coded as D, V, also see below).</u>
- *Looks up or away (LA)*. Child looks up and stares at the experimenter or looks elsewhere and does not engage with the activity at hand (is off-task). <u>It is common that while the child is off-task it also exhibits AMA. In this case, both codes should be used. If a child is actively engaged in a distraction strategy while waiting they are off-task as well, but their behaviour should be coded</u>

with a D not LA. It is possible that both the D and the LA codes are used during the same delay interval but only if a child exhibits these behaviours one after the other (e.g. plays with the dice first but then stops playing and looks away). If a genuine distraction occurs, e.g. someone walks into the room or music is suddenly heard, and the child looks up for this reason, a note should be made as follows: LA + distraction but this instance will not be counted as an instance of LA.

- *Verbalization (V).* Any verbalization by the child during a waiting interval should be coded as V. Examples might include the child making on- or off-task remarks, asking questions, requesting feedback or help, etc. If the child points to a distant object (e.g. a painting in the room) and talks about it, this behaviour should be coded as V (as opposed to D and V). The D code should only be used in conjunction with the V code when the child is talking about the experimental materials or other objects on the table in front of them that he or she touches or plays with at the time.
- *Out of seat (OS) or out of view*. Any observed instance in which the child has left his or her seat (or is out of view in case the session is videotaped). If the child suddenly stands up, for example, this behaviour, should be coded as OS.

Appendix A.8 CONSORT (Consolidated Standards of Reporting Trials) Checklist of items for reporting trials of nonpharmacologic treatments

Section	Item	Standard CONSORT Description	Extension for Nonpharmacologic Trials	Reported in section
Title and abstract	1	How participants were allocated to interventions (e.g., "random allocation," "randomized," or "randomly assigned")	In the abstract, description of the experimental treatment, comparator, care providers, centres, and blinding status	N/A
Introduction				
Background	2	Scientific background and explanation of rationale		6.2
Methods				6.3
Participants	3	Eligibility criteria for participants and the settings and locations where the data were collected	When applicable, eligibility criteria for centres and those performing the interventions	6.3.2
Interventions	4	Precise details of the interventions intended for each group and how and when they were actually administered	Precise details of both the experimental treatment and comparator	6.3.3 & 6.3.1
	4A		Description of the different components of the interventions and, when applicable, descriptions of the procedure for tailoring the interventions to individual participants	6.3.3
	4B		Details of how the interventions were standardized	6.3.3
	4C		Details of how adherence of care providers with the protocol was assessed or enhanced	6.3.3
Objectives	5	Specific objectives and hypotheses		6.3.4
Outcomes	6	Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors)		6.3.5
Sample size	7	How sample size was determined and, when applicable, explanation of any interim analyses and	When applicable, details of whether and how the clustering by care providers or centres was addressed	6.3.6

Appendices

Section	Item	Standard CONSORT Description	Extension for Nonpharmacologic Trials	Reported in section
		stopping rules		
Randomization–sequence generation	8	Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification)	When applicable, how care providers were allocated to each trial group	6.3.7
Allocation concealment	9	Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned		6.3.7
Implementation	10	Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups		6.3.7
Blinding (masking)	11A	Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment	Whether or not those administering co-interventions were blinded to group assignment	6.3.7
	11B		If blinded, method of blinding and description of the similarity of interventions†	6.3.7
Statistical methods	12	Statistical methods used to compare groups for primary outcome(s); methods for additional analyses, such as subgroup analyses and adjusted analyses	When applicable, details of whether and how the clustering by care providers or centres was addressed	6.3.8
Results				6.4
Participant flow	13	Flow of participants through each stage (a diagram is strongly recommended)---specifically, for each group, report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analysed for the primary outcome; describe deviations from study as planned, together with reasons	The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each center	6.4.1. & 6.3.2.
Implementation of intervention	New item		Details of the experimental treatment and comparator as they were implemented	6.4.1.
Recruitment	14	Dates defining the periods of recruitment and follow-up		6.3.2.
Baseline data	15	Baseline demographic and clinical characteristics of	When applicable, a description of care providers	6.4.1.

Section	Item	Standard CONSORT Description	Extension for Nonpharmacologic Trials	Reported in section
		each group	(case volume, qualification, expertise, etc.) and centres (volume) in each group	
Numbers analyzed	16	Number of participants (denominator) in each group included in each analysis and whether analysis was by "intention-to-treat"; state the results in absolute numbers when feasible (e.g., 10/20, not 50%)		6.4.1.
Outcomes and estimation	17	For each primary and secondary outcome, a summary of results for each group and the estimated effect size and its precision (e.g., 95% confidence interval)		6.4.3-6.4.6
Ancillary analyses	18	Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory		N/A
Adverse events	19	All important adverse events or side effects in each intervention group		N/A
Discussion				
Interpretation	20	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes	In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centres in each group	6.5
Generalizability	21	Generalizability (external validity) of the trial findings	Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, and care providers and centres involved in the trial	6.5
Overall evidence	22	General interpretation of the results in the context of current evidence		6.5

Appendix A.9 Information sheet and ethics forms – Schools

[school address]

[date]

Dear Mr/Mrs X.,

Re: Educational games research

We are happy to invite your school to take part in a new phase of this research.

We would also like to take this opportunity to warmly thank all the schools that have participated in earlier phases of this research for all their interest, hospitality and support.

We are currently testing a program of educational games and we would like to invite children from your school to take part. These games are designed to help improve the school readiness of children attending Year R. Please also see the attached information sheet for further details about the study.

This study has been reviewed and approved by the University of Southampton, School of Psychology Ethics Committee and is funded by a doctoral studentship to Pavlina Markomichali.

We very much hope you will agree to participate in our study. If you would like to know more about this research, or if you have any queries, please feel free to contact the project coordinator directly. We will also be contacting your school by phone or e-mail in a few days time to inquire about possible interest in the study.

Thank you in advance for all your time and your interest.

Yours sincerely

Pavlina Markomichali
Project Coordinator
MPhil/PhD Student
Developmental Brain-Behaviour Lab
School of Psychology
University of Southampton
Highfield, Southampton SO17 1BJ
e-mail: P.Markomichali@soton.ac.uk
Tel: (023) 8059 4593

Educational games study

Information sheet

UNIVERSITY OF Southampton
School of Psychology

We would like to invite your school to take part in a research study. We thought you might find the following information about the study helpful. Please do ask us if there is anything that is not clear or if there is anything you would like more information about.

What is the purpose of this study?

The purpose of this study is to test a program of educational games that is currently under development. These games are designed to help improve the school readiness of children attending Year R. The current study will be testing one particular game which aims to help children tolerate waiting situations better.

Who can take part in this study?

Children in Year R who have particular difficulties handling waiting situations, e.g. waiting for their turn, are eligible to take part. These children will be selected on the basis of their teachers' answers on a brief questionnaire screening for this sort of difficulties.

What will happen to the children if they take part?

Each child will need to complete from 2 up to 12 playing sessions. Our games are easy to learn and children will earn reward stickers and small gifts for their participation. Each playing session lasts for about 5 to 10 minutes. Only the first and the last of these sessions are expected to last slightly longer, up to 20 minutes each. All sessions will have to take place on separate days. Children will also be asked to complete one play session two to three months later. All children who take part in the study will be awarded a fun award certificate on their last session.

What will happen if children do not want to carry on with the study?

Children who volunteer to take part will be informed at the beginning of each playing session that they can stop playing the game at any time if they so wish.

What will the school need to do?

The school will need to make a quiet room or a quiet corner in a classroom available to the researcher. Teachers of children who complete the study will be asked to fill out three brief questionnaires for each child that completes all playing sessions. These questionnaires will need to be completed at every point of assessment (before the study starts, after its completion and two to three months later). Letters may need to be sent out to parents requiring opt-in or opt-out consent for the study (all relevant documentation will be provided by the researcher). Alternatively, the head teacher of your school may wish to consent to the children's participation instead.

What are the possible disadvantages and risks of taking part?

There are no disadvantages or risks involved in taking part in this study.

What are the possible benefits of taking part?

At this stage, there are no direct benefits to the children. However, as a means of compensation to the school, the study will offer Waterstones vouchers worth £25 each to each class that takes part in the study, regardless of the number of children who actually complete all the sessions. In addition, the information we will get from this study will help us improve many aspects of the program we are currently developing.

What will happen to the findings of the research study?

As explained above, the results from this study will mainly inform and improve our educational games program. It is possible that the findings from this and other related studies are presented in academic forums and also submitted for publication in academic journals. It is important to note that none of the children will be named or identified in any presentation of the findings. A copy of the summary of the research findings will be provided by the project coordinator Pavlina Markomichali once the research is finished.

What if there is a problem?

If you have any complaints, concerns, or questions about this research, please feel free to contact the project coordinator Pavlina Markomichali at the School of Psychology, University of Southampton, Southampton, SO17 1BJ (P.Markomichali@soton.ac.uk). If you wish to complain formally you can also write to the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

Will the results of this study be kept confidential?

All information collected will be held in accordance with the Data Protection Act 1988. Personal information will not be released or viewed by anyone other than the researchers involved in this project. All of the data collected will be coded so that it is anonymous and will be stored securely for ten years before it is destroyed.

Who is organizing and funding this research?

This study is organized by the University of Southampton and is funded by a doctoral studentship to Pavlina Markomichali.

Who has reviewed the study?

This study has been reviewed and approved by the University of Southampton, School of Psychology Ethics Committee.

Further information and contact details

Pavlina Markomichali
[contact details]

Educational games study
Teachers' Consent Form (v.4.)

UNIVERSITY OF Southampton
School of Psychology

Study title: Educational games study
Researcher name: Pavlina Markomichali
ERGO Study ID: 2407

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (Version 4_1-6-2012) and have had the opportunity to ask questions about the study ☐

I agree to take part in this research and I agree for my data to be used for the purpose of this study ☐

I understand my participation is voluntary and that I may withdraw my participation at any time without my legal rights being affected ☐

Name of participant ____
(Please print)

Signature ____

Date ____

This study has been reviewed and approved by University of Southampton, School of Psychology Ethics Committee. If you have any complaints, concerns, or questions about this research, please feel free to contact the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

UNIVERSITY OF Southampton
School of Psychology

Educational games study

Teachers' Debriefing Form (v.4)

Study title: Educational games study
Researcher name: Pavlina Markomichali
ERGO Study ID: 2407
RGO Ref. No:

The aim of this research is to evaluate of a program of educational games which has been designed to help very young children improve their school readiness. All the information we collected is very useful to us and it will help inform and improve many aspects of the intervention that we are planning to implement on an experimental basis in the near future.

If you would be interested in obtaining a copy of the results once the study is complete, please do not hesitate to contact me at P.Markomichali@soton.ac.uk or by phone at (02380) 594593.

If you have any complaints, concerns, or questions about this research, please feel free to contact, Chair of the Ethics Committee, Department of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

Signature ______________________________ Date ____________________

Pavlina Markomichali
Project Coordinator
PhD Research Student
Developmental Brain-Behaviour Lab
School of Psychology
University of Southampton
Highfield, Southampton SO17 1BJ
e-mail: P.Markomichali@soton.ac.uk
Tel: (02380) 594593

Appendix A.10 Information sheet and ethics forms – Parents

UNIVERSITY OF Southampton
School of Psychology

Educational games study

Information sheet for parents/guardians

We would like to invite your child to take part in a research study. We thought you might find the following information about the study helpful. Please do ask us if there is anything that is not clear or if there is anything you would like more information about.

What is the purpose of this study?

The purpose of this study is to test a program of educational games that is currently under development. These games are designed to help improve the school readiness of children attending Year R. The current study will be testing one particular game which aims to help children tolerate waiting situations better.

Why has my child been invited?

Your child has been invited to take part because they are attending Year R.

Does my child have to take part?

No, it's up to you and your child to decide. Your child will be asked if they wish to take part in the study by their teachers. If you wish for your child NOT to be asked, please return the consent form enclosed in this letter. If you do not return this form and your child agrees to take part in the study, they will still be free to withdraw at any time, without giving a reason. This will not affect the standard of care that your child receives.

What will happen to my child if they take part?

Your child will help us test a fun board game. Our games are easy to learn and your child will earn reward stickers and small toys for their participation. If your child agrees to take part, they will be asked to complete from two up to twelve playing sessions, each one lasting for 5 to 10 minutes. The first and the last of these sessions are expected to last slightly longer, up to 20 minutes each. Children will also be asked to participate in a play session two to three months later. All children who take part will also be awarded a fun award certificate on their last session.

What will happen if my child does not want to carry on with the study?

Children who volunteer to take part will be reminded at the beginning of each playing session that they can stop playing the game at any time if they so wish.

What are the possible disadvantages and risks of taking part?

There are no disadvantages or risks involved in taking part in this study.

What are the possible benefits of taking part?

At this stage, there are no direct benefits to the children. However, as a means of compensation to the school, the study will offer Waterstones vouchers worth £25 each to each class that takes part in the study, regardless of the number of children who actually complete all the sessions. In addition, the information we will get from this study will help us improve many aspects of the program we are currently developing.

What will happen to the findings of the research study?

As explained above, the results from this study will mainly inform and improve our educational games program. It is also possible that the findings from this and other related studies are presented in academic forums and also submitted for publication in academic journals. It is important to note that none of the children will be named or identified in any presentation of the findings. A copy of the summary of the research findings will be provided by the project coordinator Pavlina Markomichali once the research is finished.

What if there is a problem?

If you have any complaints, concerns, or questions about this research, please feel free to contact the project coordinator Pavlina Markomichali at the School of Psychology, University of Southampton, Southampton, SO17 1BJ (P.Markomichali@soton.ac.uk). If you wish to complain formally you can also write to the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

Will the results of this study be kept confidential?

All information collected will be held in accordance with the Data Protection Act 1988. Personal information will not be released or viewed by anyone other than the researchers involved in this project. All of the data collected will be coded so that it is anonymous and will be stored securely for ten years before it is destroyed.

Who is organizing and funding this research?

This study is organized by the University of Southampton and is funded by a doctoral studentship to Pavlina Markomichali.

Who has reviewed the study?

This study has been reviewed and approved by the University of Southampton, School of Psychology Ethics Committee.

Further information and contact details:

Pavlina Markomichali

Project Coordinator
MPhil/PhD Student
Developmental Brain-Behaviour Lab
School of Psychology
University of Southampton
Highfield, Southampton SO17 1BJ
e-mail: P.Markomichali@soton.ac.uk
Tel: (02380) 594593

Educational games study

Parental Opt-out Form (v.4.)

UNIVERSITY OF Southampton
School of Psychology

Study title: Educational games study
Researcher name: Pavlina Markomichali
ERGO Study ID: 2407

Please complete this form ONLY if you do NOT want your child to be asked to take part in the 'Educational Games Study'

If you would like to talk to the researchers about any aspect of this research before you make a decision about your child's participation, please contact Pavlina Markomichali by phone at (02380) 594593 or by e-mail at P.Markomichali@soton.ac.uk

I DO NOT give permission for my child to participate in the Educational Games Study.

Name of parent/guardian ____
(Please print)

Signature ____

Date ____

This study has been reviewed and approved by University of Southampton, School of Psychology Ethics Committee. If you have any complaints, concerns, or questions about this research, please feel free to contact the Chair of the EthicsCommittee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

UNIVERSITY OF Southampton
School of Psychology

Educational games study

Information sheet for parents/guardians

We would like to invite your child to take part in a research study. We thought you might find the following information about the study helpful. Please do ask us if there is anything that is not clear or if there is anything you would like more information about.

What is the purpose of this study?

The purpose of this study is to test a program of educational games that is currently under development. These games are designed to help improve the school readiness of children attending Year R. The current study will be testing one particular game which aims to help children tolerate waiting situations better.

Why has my child been invited?

Your child has been invited to take part because they are attending Year R.

Does my child have to take part?

No, it's up to you and your child to decide. Your child will be asked if they wish to take part in the study by their teachers. If you WISH for your child to be asked, please return the consent form enclosed in this letter. If you return this form and your child agrees to take part in the study, they will still be free to withdraw at any time, without giving a reason. This will not affect the standard of care that your child receives.

What will happen to my child if they take part?

Your child will help us test a fun board game. Our games are easy to learn and your child will earn reward stickers and small toys for their participation. If your child agrees to take part, they will be asked to complete from two up to twelve playing sessions, each one lasting for 5 to 10 minutes. The first and the last of these sessions are expected to last slightly longer, up to 20 minutes each. Children will also be asked to participate in a play session two to three months later. All children who take part will also be awarded a fun award certificate on their last session.

What will happen if my child does not want to carry on with the study?

Children who volunteer to take part will be reminded at the beginning of each playing session that they can stop playing the game at any time if they so wish.

What are the possible disadvantages and risks of taking part?

There are no disadvantages or risks involved in taking part in this study.

What are the possible benefits of taking part?

At this stage, there are no direct benefits to the children. However, as a means of compensation to the school, the study will offer Waterstones vouchers worth £25 each to each class that takes part in the study, regardless of the number of children who actually complete all the sessions. In addition, the information we will get from this study will help us improve many aspects of the program we are currently developing.

What will happen to the findings of the research study?

As explained above, the results from this study will mainly inform and improve our educational games program. It is also possible that the findings from this and other related studies are presented in academic forums and also submitted for publication in academic journals. It is important to note that none of the children will be named or identified in any presentation of the findings. A copy of the summary of the research findings will be provided by the project coordinator Pavlina Markomichali once the research is finished.

What if there is a problem?

If you have any complaints, concerns, or questions about this research, please feel free to contact the project coordinator Pavlina Markomichali at the School of Psychology, University of Southampton, Southampton, SO17 1BJ (P.Markomichali@soton.ac.uk). If you wish to complain formally you can also write to the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

Will the results of this study be kept confidential?

All information collected will be held in accordance with the Data Protection Act 1988. Personal information will not be released or viewed by anyone other than the researchers involved in this project. All of the data collected will be coded so that it is anonymous and will be stored securely for ten years before it is destroyed.

Who is organizing and funding this research?

This study is organized by the University of Southampton and is funded by a doctoral studentship to Pavlina Markomichali.

Who has reviewed the study?

This study has been reviewed and approved by the University of Southampton, School of Psychology Ethics Committee.

Further information and contact details:

Pavlina Markomichali

Project Coordinator
MPhil/PhD Student
Developmental Brain-Behaviour Lab
School of Psychology
University of Southampton
Highfield, Southampton SO17 1BJ
e-mail: P.Markomichali@soton.ac.uk
Tel: (02380) 594593

Educational games study

Parental Opt-in Form (v.4.)

Study title: Educational games study

Researcher name: Pavlina Markomichali

ERGO Study ID: 2407

Please complete this form if you WISH for your child to be asked to take part in the 'Educational Games Study'

If you would like to talk to the researchers about any aspect of this research before you make a decision about your child's participation, please contact Pavlina Markomichali by phone at (02380) 594593 or by e-mail at P.Markomichali@soton.ac.uk

I give permission for my child to participate in the Educational Games Study.

Name of parent/guardian ____
(Please print)

Signature ____

Date ____

This study has been reviewed and approved by University of Southampton, School of Psychology Ethics Committee. If you have any complaints, concerns, or questions about this research, please feel free to contact the Chair of the Ethics Committee, School of Psychology, University of Southampton, Southampton, SO17 1BJ. Phone: (02380) 594663.

Appendix A.11 WG Manual

Adapted Waiting Game training manual

[Please note the underlined sections represent updates in relation to the previous WG versions]

1.1. Equipment:

- 3 transparent colour-coded plastic discs
- colour dice
- blank sticker board (one per child)
- 2 sticker boxes, one containing plain and the other sparkly stickers
- stop-clock
- coding sheet

1.2. Reward structure:

At the beginning of the first training session, children are given a blank sticker board for collecting the stickers they win at the end of every trial. It is explained to them that the sticker boards will be kept by the experimenter (in named envelopes) for the duration of the study but will be theirs to keep after the experiment has ended. At that time they will also get a classroom certificate/award (pre-approved by the classroom teacher).

1.3. Game instructions:

Three transparent plastic discs are placed in front of the child, with one plain reward sticker placed underneath each disc. The stickers are visible to the child as the discs are transparent. Each disc has a coloured dot on it (one blue, one red and one green). The child is given a 3-colour dice to throw at the beginning of each trial.

The experimenter says: *"In this game you throw the dice to win stickers. When you throw the dice you win the sticker under the circle with the same colour on it. For example, when you throw blue, you can take the sticker under the circle with the blue dot.*

The sticker under that circle is yours, you can have the sticker and you can stick it on your sticker board BUT in this game I want you to wait until I say "Time is up!" BEFORE you can get the sticker. If you wait, then you can get a second sticker from the special box with the sparkly stickers. YOU CAN THEN KEEP BOTH STICKERS. Remember, if you do not wait, <u>or if you touch the sticker's cover</u>, you cannot pick a second sticker. Let's try!"

Two practice trials are administered at the beginning of the first training session, one where the experimenter waits until the end of the trial and picks up a second sticker and one where they experimenter does not wait and doesn't get a second sticker. In subsequent sessions, practice trials are administered only if necessary.

At the end of each trial, and depending on the outcome, the experimenter says:

"Good job/Well done/Very good waiting! You can take this sticker now and you can pick a second sticker from this box because you waited until I said "Time's up!"

"Oh dear! This sticker is yours, but you cannot pick a second sticker this time because you didn't wait for me to say time's up/you touched the cover before I said time's up."

At the start of subsequent training sessions, the experimenter says

"Now that you know how to play this game, sometimes I am going to ask you to wait for a bit longer. But if you are tired and want to stop, we can stop playing at any time."

1.4. Training protocol:

On the first training day children will be asked to wait for a total of 140 seconds (all children start on Training Level 1). On subsequent days, the total duration of the delay intervals administered will increase or decrease as shown below, depending on the children's performance on the previous day (for more details on the advancement criterion, see paragraph 1.5).

	Training Level						
Trial	0	0+	1	2	3	4	5
1	5	5	5	5	5	5	10
2	5	10	15	20	25	25	25
3	5	20	25	30	30	30	35
4	5	15	20	25	30	35	40
5	5	10	30	30	30	30	30
6	5	5	5	10	10	15	20
7	5	10	30	30	35	40	40
8	5	5	10	10	15	20	20
Total	**40**	**80**	**140**	**160**	**180**	**200**	**220**

1.5. Scoring and Advancement Criterion:

Each trial is scored as 0 = not inhibited (when the sticker cover is lifted and the sticker is actually taken), 1 = partially inhibited (when during delay there is a movement towards the cover, e.g. touching the cover or lifting the cover and touching the sticker but without making any attempt to actually take and use the sticker) and 2 = fully inhibited. The possible range of score is 0 – 16 with a high score indicating lack of impulsivity.

To avoid confusion, children will only be refused a second sticker if they have actually touched the sticker covers or the actual sticker. However, the ROMA code can still be used if a child has exhibited subtle ROMA behaviours which do not involve touching the sticker covers (such as leaning towards the stickers or their hands hovering above the stickers). In this case the trial is allowed to time out and the child still receives the second sticker, since they have not broken the touching rule. But a trial can only be scored as partially inhibited only if the child has violated the no-touching rule.

Advancement criterion: children should be moved to the next training level if they had a score of more than 11 out of 16 on the last time they trained (approximately 68% of the highest possible score), stay on the same level if they scored between 6-10 or moved down a level if their score was equal or less than 5 (or 31% of the highest possible score).

Free Waiting: At the end of every training day, the experimenter asks the children to pick a sticker that they like and wait for it for as long as they are able. To introduce the free waiting trial the experimenter says: *"Well done! Now I want you to pick your favourite sticker from the special box of stickers. Please wait as long as you can before sticking it to your sticker board. When you cannot wait any longer just let me know."* The experimenter records the number of seconds each child waits during the free waiting trial.

1.6. Observation Coding Manual

A brief behavioural checklist is completed by the experimenter during delay intervals. Behaviours falling into the behavioural categories described below are coded using a whole-interval method. This means that every time a particular behaviour occurs, the code for that category is recorded once for the duration of the whole delay interval. If different types of behaviours are observed, additional codes are added as needed as the delay interval goes on.

Restless motor activity (RMA). The child displays gross motor movements such as tapping a foot, rocking a chair back and forth, fidgeting, etc.

Reward-oriented motor activity (ROMA). Child leans towards and touches or lifts up the covers or in fact takes the reward stickers when inappropriate. Sometimes it is very hard to differentiate RMA/AMA behaviours from ROMA because the child displays, for example, aimless movement while at the same time leans towards the sticker covers. In such instances, all relevant codes are used. Note that if the ROMA code is used, then the trial is always scored as a partially inhibited (or as not inhibited, if the child actually retrieves the sticker). If the child touches the sticker covers, then the trial is terminated and the child is refused the second sticker. If other ROMA-type behaviours are exhibited, then the ROMA code is recorded (and the trial is scored as 1) but the trial is allowed to time out and the child receives the second sticker. Note that if the child hovers over or even touches the sticker cover during the last 2-3 seconds of a delay interval, in anticipation of the end of the trial, then the ROMA code should be used but the trial should not be terminated and the second sticker should not be withdrawn.

Overall Distraction Index (ODI) categories:

Aimless motor activity (AMA). The child displays aimless, waiting specific movements while waiting (touching hair, playing with clothes or jewellery, etc.). Sometimes these movements may accompany gross body movements. If this happens, both codes (RMA and AMA) are recorded. Ordinary movements such as changing body or hand positions or pointing and gesturing while speaking are not recorded as AMA, unless they are very repetitive or done in an exaggerated manner.

Distracted Attention (D). Child plays with the stickers she has won so far, the dice, the stopwatch or other objects while waiting. This code aims to capture attempts by the child to distract herself while waiting using the experimental materials. But it is possible that other objects, e.g. a toy that the child has brought along, can be used as distractors as well. But note the difference between fiddling with a bracelet (which should be coded as AMA) and taking a bracelet off to show to the experimenter and use as a talking point (which should be coded as D, V, also see below).

Looks up or away (LA). Child looks up and stares at the experimenter or looks elsewhere and does not engage with the activity at hand (is off-task). It is common that while the child is off-task it also exhibits AMA. In this case, both codes should be used. If a child is actively engaged in a distraction strategy while waiting they are off-task as well, but their behaviour should be coded with a D not LA. It is possible that both the D and the LA codes are used during the same delay interval but only if a child exhibits these behaviours one after the other (e.g. plays with the dice first but then stops playing and looks away). If a genuine distraction occurs, e.g. someone walks into the room or music is suddenly heard, and the child looks up for this reason, a note should be made as follows: LA + distraction but this instance will not be counted as an instance of LA.

Verbalization (V). Any verbalization by the child during a waiting interval should be coded as V. Examples might include the child making on- or off-task remarks, asking questions, requesting feedback or help, etc. If the child points to a distant object (e.g. a painting in the room) and talks about it, this behaviour should be coded as V (as opposed to D and V). The D code should only be used in conjunction with the V code when the child is talking about the experimental materials or other objects on the table in front of them that he or she touches or plays with at the time.

Other Vocalizations (Vo). Any instances of humming, singing or sighing can be coded as VO.

Out of seat (OS) or out of view. Any observed instance in which the child has left his or her seat (or is out of view in case the session is videotaped). If the child suddenly stands up, for example, this behaviour, should be coded as OS.

Appendix A.12 Delay Aversion Self-report Scale for Preschoolers – Trait version (DASS-P-T)

DASS-P-T. Item 1

Do you like waiting for things? Like when you have to wait in the queue before lunch? (Response Options: Yes/No/I do not know)

DASS-P-T. Item 2

When you have to wait for things, can you show me how you feel? (Response Options: happy, sad or you do not mind very much)

DASS-P-T. Item 3

3. When you have to wait for things, can you show me how hard you have to try? (Response Options: not so hard, a bit hard, very hard)

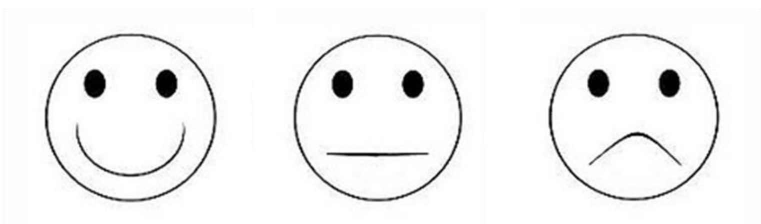
Appendix A.13 Delay Aversion Self-report Scale for Preschoolers – State version (DASS-P-S)

DASS-P-S. Item 1

Did you like the waiting today? When you were waiting before you could get your stickers (Response Options: Yes/No/I do not know)

DASS-P-S. Item 2

How did you feel during the waiting today? (Response Options: happy, sad or you did not mind it very much)



DASS-P-S. Item 3

How hard did you find the waiting today? (Response Options: not so hard, a bit hard, very hard)

Bibliography

Abikoff, H. (1991). Cognitive training in ADHD children: Less to it than meets the eye. *Journal of Learning Disabilities, 24*, 205-209.

Alberts, E., & van der Meere, J. (1992). Observations of hyperactive behavior during vigilance. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(8), 1355-1364.

American Academy of Pediatrics. (2011). ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. *Pediatrics, 128*(5), 1007-1022.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anokhin, A. P., Golosheykin, S., Grant, J. D., & Heath, A. C. (2011). Heritability of Delay Discounting in Adolescence: A Longitudinal Twin Study. *Behavior Genetics, 41*(2), 175-183.

Antrop, I., Roeyers, H., Van Oost, P., & Buysse, A. (2000). Stimulation seeking and hyperactivity in children with ADHD. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(2), 225-231.

Antrop, I., Stock, P., Verte, S., Wiersema, J. R., Baeyens, D., & Roeyers, H. (2006). ADHD and delay aversion: the influence of non-temporal stimulation on choice for delayed rewards. *Journal of Child Psychology and Psychiatry, 47*(11), 1152-1158.

Antshel, K. M., & Barkley, R. (2008). Psychosocial interventions in attention deficit hyperactivity disorder. *Child and Adolescent Psychiatric Clinics of North America, 17*(2), 421.

Banaschewski, T., Coghill, D., Santosh, P., Zuddas, A., Asherson, P., Buitelaar, J., et al. (2006). Long-acting medications for the hyperkinetic disorders - A systematic review and European treatment guideline. *European Child & Adolescent Psychiatry, 15*(8), 476-495.

Barkley, R. A. (1997a). *ADHD and the Nature of Self-control*. New York The Guilford Press.

Barkley, R. A. (1997b). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94.

Berry, M. S., Sweeney, M. M., Morath, J., Odum, A. L., & Jordan, K. E. (2014). The Nature of Impulsivity: Visual Exposure to Natural Environments Decreases Impulsive Decision-Making in a Delay Discounting Task. *PLoS ONE, 9*(5), e97915.

Bibliography

Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., & Baxter, C. (2011). Remember the Future: Working Memory Training Decreases Delay Discounting Among Stimulant Addicts. *Biological Psychiatry, 69*(3), 260-265.

Binder, L. M., Dixon, M. R., & Ghezzi, P. M. (2000). A procedure to teach self-control to children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 33*(2), 233-237.

Bitsakou, P. (2007). *Dissecting the neuropsychological structure of ADHD: inhibitory control and delay aversion.* University of Southampton, Southampton.

Bitsakou, P., Antrop, I., Wiersema, J. R., & Sonuga-Barke, E. J. S. (2006). Probing the limits of delay intolerance: Preliminary young adult data from the Delay Frustration Task (DeFT). [Article]. *Journal of Neuroscience Methods, 151*(1), 38-44.

Bitsakou, P., Psychogiou, L., Thompson, M., & Sonuga-Barke, E. J. S. (2009). Delay Aversion in Attention Deficit/Hyperactivity Disorder: An empirical investigation of the broader phenotype. *Neuropsychologia, 47*(2), 446-456.

Bjork, J. M., Knutson, B., Fong, G. W., Caggiano, D. M., Bennett, S. M., & Hommer, D. W. (2004). Incentive-elicited brain activation in adolescents: Similarities and differences from young adults. *Journal of Neuroscience, 24*(8), 1793-1802.

Black, A. C., & Rosen, M. I. (2011). A money management-based substance use treatment increases valuation of future rewards. *Addictive Behaviors, 36*(1-2), 125-128.

Blair, C. (2002). School readiness - Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*(2), 111-127.

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: The promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology, 20*(3), 899-911.

Blair, C., & Peters Razza, R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647-663.

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., Ravaud, P., & Group, C. (2008). Methods and Processes of the CONSORT Group: Example of an Extension for Trials Assessing Nonpharmacologic Treatments. *Annals of Internal Medicine, 148*(4), W60-W66.

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., Ravaud, P., & Grp, C. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Annals of Internal Medicine, 148*(4), 295-309.

Calkins, S. D. (2007). The emergence of self-regulation: Biological and behavioural control mechanisms supporting toddler competencies. In C. A. Brownell & C. B. Kopp (Eds.), *Socio-emotional development in the toddler years: Transitions and tranformations* (pp. 261-284). New York, NY: Guildford Press.

Campbell, S. B. (2002). *Behavior problems in preschool children: clinical and developmental issues* (2nd ed.). New York, NY: Guilford Press.

Campbell, S. B., Pierce, E. W., March, C. L., Ewing, L. J., & Szumowski, E. K. (1994). Hard-to-manage preschool boys-symptomatic behavior across contexts and time. *Child Development, 65*(3), 836-851.

Campbell, S. B., Szumowski, E. K., Ewing, L. J., Gluck, D. S., & Breaux, A. M. (1982). A multi-dimensional assessment of parent-identified behavior problem toddlers. *Journal of Abnormal Child Psychology, 10*, 569-592.

Carlson, S. M. (2005). Developmentally Sensitive Measures of Executive Function in Preschool Children. *Developmental Neuropsychology, 28*(2), 595 - 616.

Cartwright, K. B. (2012). Insights From Cognitive Neuroscience: The Importance of Executive Function for Early Reading Development and Education. *Early Education and Development, 23*(1), 24-36.

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: beyond executive dysfunction. *Trends in Cognitive Sciences, 10*(3), 117-123.

Chang, Y.-K., Tsai, Y.-J., Chen, T.-T., & Hung, T.-M. (2013). The impacts of coordinative exercise on executive function in kindergarten children: an ERP study. *Experimental Brain Research, 225*(2), 187-196.

Charach, A., Carson, P., Fox, S., Ali, M. U., Beckett, J., & Lim, C. G. (2013). Interventions for Preschool Children at High Risk for ADHD: A Comparative Effectiveness Review. *Pediatrics, 131*(5), E1584-E1604.

Chorpita, B. F., Yim, L. M., Donkervoet, J. C., Arensdorf, A., Amundsen, M. J., McGee, C., et al. (2002). Toward large-scale implementation of empirically supported treatments for children: A review and observations by the Hawaii empirical basis to services task force. *Clinical Psychology-Science and Practice, 9*(2), 165-190.

Chronis, A. M., Jones, H. A., & Raggi, V. L. (2006). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. [Review]. *Clinical Psychology Review, 26*(4), 486-502.

Cicchetti, D., & Toth, S. L. (2006). Building bridges and crossing them: Translational research in developmental psychopathology. *Development and Psychopathology, 18*(3), 619-622.

Bibliography

Clare, S., Helps, S., & Sonuga-Barke, E. J. (2010). The quick delay questionnaire: a measure of delay aversion and discounting in adults. *Attention deficit and hyperactivity disorders, 2*(1), 43-48.

Coghill, D., Seth, S., & Matthews, K. (2014). A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: advancing beyond the three-pathway models. *Psychological Medicine, 44*(9), 1989-2001.

Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual Research Review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders - implications of recent empirical study. *Journal of Child Psychology and Psychiatry, 53*(5), 469-489.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Converse, P. D., & DeShon, R. P. (2009). A tale of two tasks: Reversing the self-regulatory resource depletion effect. *Journal of Applied Psychology, 94*(5), 1318-1324.

Cortese, S., Ferrin, M., Brandeis, D., Buitelaar, J., Daley, D., Dittmann, R. W., et al. (in preparation). Computerized cognitive training for children and adolescents with ADHD: a comprehensive meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials.

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis, 34*(1), 101-122.

Cubillo, A., Halari, R., Smith, A., Taylor, E., & Rubia, K. (2012). A review of fronto-striatal and fronto-cortical brain abnormalities in children and adults with Attention Deficit Hyperactivity Disorder (ADHD) and new evidence for dysfunction in adults with ADHD during motivation and attention. *Cortex, 48*(2), 194-215.

Curry, S. H. (2008). Translational science: past, present, and future. *Biotechniques, 44*(2), II-VIII.

Dalen, L., Sonuga-Barke, E. J. S., Hall, M., & Remington, B. (2004). Inhibitory deficits, delay aversion and preschool AD/HD: implications for the dual pathway model. *Neural Plasticity, 11*(1-2), 1-11.

Daley, D., Jones, K., Hutchings, J., & Thompson, M. (2009). Attention deficit hyperactivity disorder in pre-school children: current findings, recommended interventions and future directions. *Child Care Health and Development, 35*(6), 754-766.

Daly, B. P., Creed, T., Xanthopoulos, M., & Brown, R. T. (2007). Psychosocial treatments for children with attention deficit/hyperactivity disorder. *Neuropsychology Review, 17*(1), 73-89.

Demurie, E., Roeyers, H., Baeyens, D., & Sonuga-Barke, E. (2012). Temporal discounting of monetary rewards in children and adolescents with ADHD and autism spectrum disorders. *Developmental Science, 15*(6), 791-800.

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. Craik (Eds.), *Lifespan Cognition: Mechanisms of Change* (pp. 70-95). NY: Oxford University Press.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). The Early Years: Preschool Program Improves Cognitive Control. *Science, 318*(5855), 1387-1388.

Diamond, A., & Lee, K. (2011). Interventions Shown to Aid Executive Function Development in Children 4 to 12 Years Old. *Science, 333*(6045), 959-964.

Dias, T. G. C., Wilson, V. B., Bathula, D. R., Iyer, S. P., Mills, K. L., Thurlow, B. L., et al. (2013). Reward circuit connectivity relates to delay discounting in children with attention-deficit/hyperactivity disorder. *European Neuropsychopharmacology, 23*(1), 33-45.

Dichter, G. S., Damiano, C. A., & Allen, J. A. (2012). Reward circuitry dysfunction in psychiatric and neurodevelopmental disorders and genetic syndromes: animal models and clinical findings. *Journal of Neurodevelopmental Disorders, 4*.

Dixon, M. R., Hayes, L. J., Binder, L. M., Manthey, S., Sigman, C., & Zdanowski, D. M. (1998). Using a self-control training procedure to increase appropriate behavior. *Journal of Applied Behavior Analysis, 31*(2), 203.

Dixon, M. R., & Tibbetts, P. A. (2009). The effects of choice on self-control. *Journal of Applied Behavior Analysis, 42*(2), 243-252.

Douglas, V. I., & Parry, P. A. (1983). Effects of reward on delayed reaction time task performance of hyperactive children. *Journal of Abnormal Child Psychology, 11*(2), 313-326.

Douglas, V. I., & Parry, P. A. (1994). Effects of reward and nonreward on frustration and attention in attention-deficit disorder. *Journal of Abnormal Child Psychology, 22*(3), 281-302.

Dowsett, S. M., & Livesey, D. J. (2000). The development of inhibitory control in preschool children: Effects of "executive skills" training. *Developmental Psychobiology, 36*(2), 161-174.

Duku, E., & Vaillancourt, T. (2013). Validation of the BRIEF-P in a sample of Canadian preschool children. *Child Neuropsychology*, 1-14.

DuPaul, G. J. (2007). School-based interventions for students with attention deficit hyperactivity disorder: Current status and future directions. *School Psychology Review, 36*(2), 183-194.

Bibliography

Durston, S., van Belle, J., & de Zeeuw, P. (2011). Differentiating Frontostriatal and Fronto-Cerebellar Circuits in Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry, 69*(12), 1178-1184.

Edel, M.-A., Enzi, B., Witthaus, H., Tegenthoff, M., Peters, S., Juckel, G., et al. (2013). Differential reward processing in subtypes of adult attention deficit hyperactivity disorder. *Journal of Psychiatric Research, 47*(3), 350-356.

Edmunds, M. R., Barry, R. J., & Denniston, A. K. (2013). Readability Assessment of Online Ophthalmic Patient Information. *Journal of the American Medical Association Ophthalmology, 131*(12), 1610-1616.

Eigsti, I. M., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., et al. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychological Science, 17*(6), 478-484.

Espinet, S. D., Anderson, J. E., & Zelazo, P. D. (2013). Reflection training improves executive function in preschool-age children: Behavioral and neural effects. *Developmental Cognitive Neuroscience, 4*, 3-15.

Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America, 109*(17), 6769-6774.

Frederick, S., & Loewenstein, G. (2008). Conflicting motives in evaluations of sequences. *Journal of Risk and Uncertainty, 37*(2-3), 221-235.

Ghuman, J. K. (2008). Psychopharmacological and Other Treatments in Preschool Children with Attention-Deficit/Hyperactivity Disorder: Current Evidence and Practice. *Journal of Child and Adolescent Psychopharmacology, 18*, 413-447.

Gillberg, C., Gillberg, I. C., Rasmussen, P., Kadesjo, B., Soderstrom, H., Rastam, M., et al. (2004). Co-existing disorders in ADHD - implications for diagnosis and intervention. *European Child & Adolescent Psychiatry, 13*(Suppl. 1), I80-I92.

Gleason, M. M., Egger, H. L., Emslie, G. J., Greenhill, L. L., Kowatch, R. A., Lieberman, A. F., et al. (2007). Psychopharmacological treatment for very young children: Contexts and guidelines. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 1532-1572.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1337-1345.

Graham, J., Banaschewski, T., Buitelaar, J., Coghill, D., Danckaerts, M., Dittmann, R. W., et al. (2011). European guidelines on managing adverse effects of medication for ADHD. [Article]. *European Child & Adolescent Psychiatry, 20*(1), 17-37.

Graziano, P. A., Reavis, R. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation in children's early academic success. *Journal of School Psychology, 45*(1), 3-19.

Grosch, J., & Neuringer, A. (1981). Self-control in pigeons under the Mischel paradigm. *Journal of the Experimental Analysis of Behavior, 35*(1), 3-21.

Guderjahn, L., Gold, A., Stadler, G., & Gawrilow, C. (2013). Self-regulation strategies support children with ADHD to overcome symptom-related behavior in the classroom. *Attention deficit and hyperactivity disorders, 5*(4), 397-407.

Guo, Y., Logan, H. L., Glueck, D. H., & Muller, K. E. (2013). Selecting a sample size for studies with repeated measures. *BMC Medical Research Methodology, 13*, 100.

Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in clinical research, 2*(3), 109-112.

Haenlein, M., & Caul, W. F. (1987). Attention-deficit disorder with hyperactivity - a specific hypothesis of reward dysfunction. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(3), 356-362.

Halperin, J. M., Bedard, A.-C. V., & Curchack-Lichtin, J. T. (2012). Preventive Interventions for ADHD: A Neurodevelopmental Perspective. *Neurotherapeutics, 9*(3), 531-541.

Halperin, J. M., Marks, D. J., Bedard, A.-C. V., Chacko, A., Curchack, J. T., Yoon, C. A., et al. (2012). Training Executive, Attention, and Motor Skills: A Proof-of-Concept Study in Preschool Children With ADHD. *Journal of Attention Disorders*.

Halperin, J. M., Marks, D. J., Bedard, A.-C. V., Chacko, A., Curchack, J. T., Yoon, C. A., et al. (2013). Training Executive, Attention, and Motor Skills: A Proof-of-Concept Study in Preschool Children With ADHD. *Journal of Attention Disorders, 17*(8), 711-721.

Harpin, V. A. (2005). The effect of ADHD on the life of an individual, their family, and community from preschool to adult life *Archives of Disease in Childhood, 90*(Suppl. 1), I2-I7

Healey, D. M., & Halperin, J. M. (2014). Enhancing Neurobehavioral Gains with the Aid of Games and Exercise (ENGAGE): Initial open trial of a novel early intervention fostering the development of preschoolers' self-regulation. *Child Neuropsychology*, 1-16.

Huang, J., Yang, B.-r., Zou, X.-b., Jing, J., Pen, G., McAlonan, G. M., et al. (2012). Temporal processing impairment in children with attention-deficit-hyperactivity disorder. *Research in Developmental Disabilities, 33*(2), 538-548.

Jelicic, H., Phelps, E., & Lerner, R. A. (2009). Use of Missing Data Methods in Longitudinal Studies: The Persistence of Bad Practices in Developmental Psychology. *Developmental Psychology, 45*(4), 1195-1199.

Jensen, P. S., Arnold, L. E., Swanson, J. M., Vitiello, B., Abikoff, H. B., Greenhill, L. L., et al. (2007). 3-year follow-up of the NIMH MTA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(8), 989-1002.

Jimura, K., Chushak, M. S., & Braver, T. S. (2013). Impulsivity and Self-Control during Intertemporal Decision Making Linked to the Neural Dynamics of Reward Value Representation. *Journal of Neuroscience, 33*(1), 344-357.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77*(2), 129-146.

Jones, K., Daley, D., Hutchings, J., Bywater, T., & Eames, C. (2008). Efficacy of the Incredible Years Programme as an early intervention for children with conduct problems and ADHD: long-term follow-up. *Child Care Health and Development, 34*(3), 380-390.

Kanfer, F. H. (1985). Target selction for clinical change programs. *Behavioural Analysis, 7*, 7-20.

Kelly, C., Sonuga-Barke, E., Scheres, A., & Castellanos, F. (2007). Functional neuroimaging of the reward and motivational pathway in ADHD. In M. Bellgrove, M. Fitzgerald & M. Gill (Eds.), *The handbook of Attention Deficit Hyperactivity Disorder*: John Wiley & Sons, Ltd.

Klein, R. G., & Mannuzza, S. (1991). Long-term outcome of hyperactive children: a review. [Article]. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*(3), 383-387.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14*(7), 317-324.

Klingberg, T., Fernell, E., Olesen, P., Johnson, M., Gustafsson, P., Dahlström, K., et al. (2005). Computerized Training of Working Memory in Children with ADHD – a Randomized, Controlled, Trial. *Journal of American Academy of Child and Adolescent Psychiatry, 44*(2), 177-186.

Kloo, D., & Perner, J. (2003). Training transfer between card sorting and false belief understanding: Helping children apply conflicting descriptions. *Child Development, 74*(6), 1823-1839.

Knight, L. A., Rooney, M., & Chronis-Tuscano, A. (2008). Psychosocial treatments for attention-deficit/hyperactivity disorder. *Current Psychiatry Reports, 10*, 412-418.

Koerting, J., Smith, E., Knowles, M. M., Latter, S., Elsey, H., McCann, D. C., et al. (2013). Barriers to, and facilitators of, parenting programmes for childhood behaviour problems: a qualitative synthesis of studies of parents' and professionals' perceptions. *European Child & Adolescent Psychiatry, 22*(11), 653-670.

Kollins, S., Greenhill, L., Swanson, J., Wigal, S., Abikoff, H., McCracken, J., et al. (2006). Rationale, Design, and Methods of the Preschool ADHD Treatment Study (PATS). *Journal of the American Academy of Child & Adolescent Psychiatry, 45*(11), 1275-1283.

Kratochvil, C. J., Greenhill, L. L., March, J. S., Burke, W. J., & Vaughan, B. S. (2004). The Role of Stimulants in the Treatment of Preschool Children with Attention-Deficit Hyperactivity Disorder. *CNS Drugs, 18*(14), 957-966.

Kross, E., Ayduk, O., & Mischel, W. (2005). When asking "why" does not hurt - Distinguishing rumination from reflective processing of negative emotions. *Psychological Science, 16*(9), 709-715.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Kuntsi, J., Stevenson, J., Oosterlaan, J., & Sonuga-Barke, E. J. S. (2001). Test-retest reliability of a new delay aversion task and executive function measures. *British Journal of Developmental Psychology, 19*(3), 339-348.

LaForett, D. R. (2008). Psychosocial treatments for preschool-aged children with Attention-Deficit Hyperactivity Disorder. *Developmental Disabilities Research Reviews, 14*(4), 300-310.

Lahey, B. B., Pelham, W. E., Loney, J., Kipp, H., Ehrhardt, A., Lee, S. S., et al. (2004). Three-year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4-6 years of age. *American Journal of Psychiatry, 161*(11), 2014-2020.

Lahey, B. B., Pelham, W. E., Stein, M. A., Loney, J., Trapani, C., Nugent, K., et al. (1998). Validity of DSM-IV attention-deficit/hyperactivity disorder for younger children. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*(7), 695-702.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lee, S. I., Schachar, R. J., Chen, S. X., Ornstein, T. J., Charach, A., Barr, C., et al. (2008). Predictive validity of DSM-IV and ICD-10 criteria for ADHD and hyperkinetic disorder. *Journal of Child Psychology and Psychiatry, 49*(1), 70-78.

Lillard, A., & Else-Quest, N. (2006). The early years: Evaluating Montessori education. *Science, 313*(5795), 1893-1894.

Loe, I. M., & Feldman, H. M. (2007). Academic and educational outcomes of children with ADHD. *Journal of Pediatric Psychology, 32*(6), 643-654.

Logue, A. W. (1995). *Self-control: waiting until tomorrow for what you want today*. Englewood Cliffs, NJ: Prentice Hall.

Luman, M., Tripp, G., & Scheres, A. (2010). Identifying the neurobiology of altered reinforcement sensitivity in ADHD: A review and research agenda. *Neuroscience and Biobehavioral Reviews, 34*(5), 744-754.

Mahone, E. M., & Hoffman, J. (2007). Behavior rating of executive function among preschoolers with ADHD. *Clinical Neuropsychologist, 21*(4), 569-586.

Mannuzza, S., Klein, R. G., Bessler, A., Malloy, P., & LaPadula, M. (1998). Adult psychiatric status of hyperactive boys grown up. *American Journal of Psychiatry, 155*(4), 493-498.

Marco, R., Miranda, A., Schlotz, W., Melia, A., Mulligan, A., Müller, U., et al. (2009). Delay and reward choice in ADHD: An experimental test of the role of delay aversion. *Neuropsychology, 23*(3), 367-380.

Markomichali, P., Donnelly, N., & Sonuga Barke, E. J. S. (2009). Cognitive training for attention, inhibition and working memory deficits: A potential treatment for ADHD? *Advances in ADHD, 3*(3), 89-98.

Mazur, J. E. (2005). Effects of reinforcer probability, delay, and response requirements on the choices of rats and pigeons: possible species differences. *Journal of the Experimental Analysis of Behavior, 83*(3), 263-279.

McClelland, M. M., & Cameron, C. E. (2011). Self-Regulation and Academic Achievement in Elementary School Children. In R. M. Lerner, J. V. Lerner, E. P. Bowers, S. LewinBizan, S. Gestsdottir & J. B. Urban (Eds.), *Thriving in Childhood and Adolescence: The Role of Self-Regulation Processes* (Vol. 133, pp. 29-44).

McClelland, M. M., & Cameron, C. E. (2012). Self-Regulation in Early Childhood: Improving Conceptual Clarity and Developing Ecologically Valid Measures. *Child Development Perspectives, 6*(2), 136-142.

McGoey, K. E., Eckert, T. L., & DuPaul, G. J. (2002). Early Intervention for Preschool-Age Children with ADHD: A Literature Review. *Journal of Emotional and Behavioral Disorders, 10*, 14-28.

Melby-Lervag, M., & Hulme, C. (2013). Is Working Memory Training Effective? A Meta-Analytic Review. *Developmental Psychology, 49*(2), 270-291.

Miklowitz, D. J., & Cicchetti, D. (2006). Toward a life span developmental psychopathology perspective on bipolar disorder. *Development and Psychopathology, 18*(4), 935-938.

Mischel, W., & Ebbesen, E. B. (1970). Attention in delay of gratification. *Journal of Personality and Social Psychology, 16*(2), 329-&.

Mischel, W., & Metzner, R. (1962). Preference for delayed reward as a function of age, intelligence, and length of delay interval. *The Journal of Abnormal and Social Psychology, 64*(6), 425-431.

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology, 54*(4), 687-696.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244*(4907), 933-938.

Mischel, W., Zeiss, A. R., & Ebbesen, E. B. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology, 21*(2), 204-218.

Molina, B. S. G., Hinshaw, S. P., Arnold, L. E., Swanson, J. M., Pelham, W. E., Hechtman, L., et al. (2013). Adolescent Substance Use in the Multimodal Treatment Study of Attention-Deficit/Hyperactivity Disorder (ADHD) (MTA) as a Function of Childhood ADHD, Random Assignment to Childhood Treatments, and Subsequent Medication. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(3), 250-263.

Moore, B., Mischel, W., & Zeiss, A. (1976). Comparative effects of reward stimulus and its cognitive representation in voluntary delay. *Journal of Personality and Social Psychology, 34*(3), 419-424.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125.

MTA Cooperative Group. (1999). A 14-Month Randomized Clinical Trial of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder. *Archives of General Psychiatry, 56*(12), 1073-1086.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*(2), 247-259.

Neef, N. A., Bicard, D. F., & Endo, S. (2001). Assessment of impulsivity and the development of self-control in students with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 34*(4), 397-408.

Neubauer, A., Gawrilow, C., & Hasselhorn, M. (2012). The Watch-and-Wait Task: On the reliability and validity of a new method of assessing self-control in preschool children. *Learning and Individual Differences, 22*(6), 770-777.

Neuenschwander, R., Roethlisberger, M., Cimeli, P., & Roebers, C. M. (2012). How do different aspects of self-regulation predict successful adaptation to school? *Journal of Experimental Child Psychology, 113*(3), 353-371.

Nigg, J. T., & Casey, B. J. (2005). An integrative theory of attention-deficit/hyperactivity disorder based on the cognitive and affective neurosciences. *Development and Psychopathology, 17*(3), 785-806.

Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry, 57*(11), 1224-1230.

Nijmeijer, J. S., Minderaa, R. B., Buitelaar, J. K., Mulligan, A., Hartman, C. A., & Hoekstra, P. J. (2008). Attention-deficit/hyperactivity disorder and social dysfunctioning. *Clinical Psychology Review, 28*(4), 692-708.

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*(3), 427-439.

Owens, J. S., Goldfine, M. E., Evangelista, N. M., Hoza, B., & Kaiser, N. M. (2007). A critical review of self-perceptions and the positive illusory bias in children with ADHD. *Clinical Child and Family Psychology Review, 10*(4), 335-351.

Paloyelis, Y., Mehta, M. A., Faraone, S. V., Asherson, P., & Kuntsi, J. (2012). Striatal Sensitivity During Reward Processing in Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(7), 722-732.

Pauli-Pott, U., Dalir, S., Mingebach, T., Roller, A., & Becker, K. (2013). Do different ADHD-related etiological risks involve specific neuropsychological pathways? An analysis of mediation processes by inhibitory control and delay aversion. *Journal of Child Psychology and Psychiatry, 54*(7), 800-809.

Peake, P. K., Hebl, M., & Mischel, W. (2002). Strategic attention deployment for delay of gratification in working and waiting situations. *Developmental Psychology, 38*(2), 313-326.

Pelham, W. E., & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology, 37*(1), 184-214.

Persson, J., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2007). Cognitive fatigue of executive processes: Interaction between interference resolution tasks. [Article]. *Neuropsychologia, 45*(7), 1571-1579.

Pfiffner, L. J., Villodas, M., Kaiser, N., Rooney, M., & McBurnett, K. (2013). Educational Outcomes of a Collaborative School-Home Behavioral Intervention for ADHD. *School Psychology Quarterly, 28*(1), 25-36.

Platzman, K. A., Stoy, M. R., Brown, R. T., Coles, C. D., Smith, I. E., & Falek, A. (1992). Review of observational methods in attention deficit hyperactivity disorder (ADHD): Implications for diagnosis. *School Psychology Quarterly, 7*(3), 155-177.

Plichta, M. M., Vasic, N., Wolf, R. C., Lesch, K.-P., Brummer, D., Jacob, C., et al. (2009). Neural Hyporesponsiveness and Hyperresponsiveness During Immediate and Delayed Reward Processing in Adult Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry, 65*(1), 7-14.

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *American Journal of Psychiatry, 164*(6), 942-948.

Prelec, D., & Loewenstein, G. (1991). Decision-making over time and under uncertainty - a common approach. *Management Science, 37*(7), 770-786.

Quay, H. C. (1997). Inhibition and attention deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 25*(1), 7-13.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardine, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly, 23*(1), 10-26.

Reid, R., Trout, A. L., & Schartz, M. (2005). Self-regulation interventions for children with attention deficit/hyperactivity disorder. *Exceptional Children, 71*(4), 361-377.

Riddle, M. A., Yershova, K., Lazzaretto, D., Paykina, N., Yenokyan, G., Greenhill, L., et al. (2013). The Preschool Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS) 6-Year Follow-Up. *Journal of the American Academy of Child and Adolescent Psychiatry, 52*(3), 264-278.

Riggs, N. R., Greenberg, M. T., Kusche, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS curriculum. *Prevention Science, 7*(1), 91-102.

Roethlisberger, M., Neuenschwander, R., Cimeli, P., Michel, E., & Roebers, C. M. (2012). Improving executive functions in 5- and 6-year-olds: Evaluation of a small group intervention in prekindergarten and kindergarten children. *Infant and Child Development, 21*(4), 411-429.

Rohde, L. A., Barbosa, G., Polanczyk, G., Eizirik, M., Rasmussen, E. R., Neuman, R. J., et al. (2001). Factor and latent class analysis of DSM-IV ADHD symptoms in a school sample of Brazilian adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(6), 711-718.

Root, R. W., & Resnick, R. J. (2003). An update on the diagnosis and treatment of attention-deficit/hyperactivity disorder in children. *Professional Psychology-Research and Practice, 34*(1), 34-41.

Rubia, K., Noorloos, J., Smith, A., Gunning, B., & Sergeant, J. (2003). Motor timing deficits in community and clinical boys with hyperactive behavior: The effect of methylphenidate on motor timing. *Journal of Abnormal Child Psychology, 31*(3), 301-313.

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia, 42*(8), 1029-1040.

Rueda, M. R., Rothbart, M. K., McCandliss, B. D., Saccomanno, L., & Posner, M. I. (2005). Training, maturation, and genetic influences on the development of executive attention. *Proceedings of the National Academy of Sciences of the United States of America, 102*(41), 14931-14936.

Rutherford, A. (2011). *ANOVA and ANCOVA: a GLM approach* (2nd ed. ed.). Oxford: Wiley.

Rutledge, K. J., van den Bos, W., McClure, S. M., & Schweitzer, J. B. (2012). Training Cognition in ADHD: Current Findings, Borrowed Concepts, and Future Directions. *Neurotherapeutics, 9*(3), 542-558.

Sagvolden, T., Aase, H., Zeiner, P., & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research, 94*(1), 61-71.

Sagvolden, T., Johansen, E. B., Aase, H., & Russell, V. A. (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behavioral and Brain Sciences, 28*(3), 397.

Sayal, K., Taylor, E., Beecham, J., & Byrne, P. (2002). Pathways to care in children at risk of attention-deficit hyperactivity disorder. *The British Journal of Psychiatry, 181*(1), 43-48.

Scheres, A., Dijkstra, M., Ainslie, E., Balkan, J., Reynolds, B., Sonuga-Barke, E., et al. (2006). Temporal and probabilistic discounting of rewards in children and adolescents: Effects of age and ADHD symptoms. *Neuropsychologia, 44*(11), 2092-2103.

Scheres, A., Milham, M. P., Knutson, B., & Castellanos, F. X. (2007). Ventral striatal hyporesponsiveness during reward anticipation in attention-deficit/hyperactivity disorder. *Biological Psychiatry, 61*(5), 720-724.

Scheres, A., Tontsch, C., & Thoeny, A. L. (2013). Steep temporal reward discounting in ADHD-Combined type: Acting upon feelings. *Psychiatry Research, 209*(2), 207-213.

Schunk, D. H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review, 20*(4), 463-467.

Schweitzer, J. B., & Sulzer-Azaroff, B. (1988). Self-control: teaching tolerance for delay in impulsive children. *Journal of Experimental Analysis of Behavior, 50*, 173-186.

Schweitzer, J. B., & Sulzer-Azaroff, B. (1995). Self-control in boys with attention-deficit hyperactivity disorder - effects of added stimulation and time. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*(4), 671-686.

Sergeant, J. (2000). The cognitive-energetic model: an empirical approach to Attention-Deficit Hyperactivity Disorder. *Neuroscience and Biobehavioral Reviews 24*(1), 7-12.

Shamosh, N. A., DeYoung, C. G., Green, A. E., Reis, D. L., Johnson, M. R., Conway, A. R. A., et al. (2008). Individual Differences in Delay Discounting Relation to Intelligence, Working Memory, and Anterior Prefrontal Cortex. *Psychological Science, 19*(9), 904-911.

Sherman, E. M. S., & Brooks, B. L. (2010). Behavior Rating Inventory of Executive Function – Preschool Version (BRIEF-P): Test Review and Clinical Guidelines for Use. *Child Neuropsychology, 16*(5), 503-519.

Shiels, K., & Hawk, L. W., Jr. (2010). Self-regulation in ADHD: The role of error processing. *Clinical Psychology Review, 30*(8), 951-961.

Shiels, K., Hawk, L. W., Jr., Reynolds, B., Mazzullo, R. J., Rhodes, J. D., Pelham, W. E., Jr., et al. (2009). Effects of Methylphenidate on Discounting of Delayed Rewards in Attention Deficit/Hyperactivity Disorder. *Experimental and Clinical Psychopharmacology, 17*(5), 291-301.

Shinaver, C. S., 3rd, Entwistle, P. C., & Soderqvist, S. (2014). Cogmed WM Training: Reviewing the Reviews. *Applied neuropsychology. Child, 3*(3), 163-172.

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is Working Memory Training Effective? *Psychological Bulletin, 138*(4), 628-654.

Sjowall, D., Roth, L., Lindqvist, S., & Thorell, L. B. (2013). Multiple deficits in ADHD: executive dysfunction, delay aversion, reaction time variability, and emotional deficits. *Journal of Child Psychology and Psychiatry, 54*(6), 619-627.

Sjowall, D., & Thorell, L. B. (2014). Functional Impairments in Attention Deficit Hyperactivity Disorder: The Mediating Role of Neuropsychological Functioning. *Developmental Neuropsychology, 39*(3), 187-204.

Skinner, B. F. (1948). *Walden two*. Indianapolis: Hackett Publishing Company, Inc.

Sohlberg, M. M., & Mateer, C. A. (2001). Improving attention and managing attentional problems. Adapting rehabilitation techniques to adults with ADD. *Annals of the New York Academy of Sciences, 931*, 359-375.

Solanto, M. V., Abikoff, H., Sonuga-Barke, E. J. S., Schachar, R., Logan, G. D., Wigal, T., et al. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: A supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology, 29*(3), 215-228.

Sonuga-Barke, E. J. S. (2002). Psychological heterogeneity in AD/HD: A dual pathway model of behaviour and cognition. *Behavioural Brain Research, 130*(1-2), 29-36.

Sonuga-Barke, E. J. S. (2003). The dual pathway model of AD/HD: an elaboration of neuro-developmental characteristics. *Neuroscience and Biobehavioral Reviews, 27*(7), 593-604.

Sonuga-Barke, E. J. S. (2004). On the Reorganization of Incentive Structure to Promote Delay Tolerance: A Therapeutic Possibility for AD/HD? *Neural Plasticity, 11*(1-2), 23-28.

Sonuga-Barke, E. J. S. (2005). Causal models of attention-deficit/hyperactivity disorder: From common simple deficits to multiple developmental pathways. *Biological Psychiatry, 57*(11), 1231-1238.

Sonuga-Barke, E. J. S. (2013). Attention-Deficit/Hyperactivity Disorder: Toward a developmental synthesis In P. D. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology: Volume 1 Body and Mind* (Vol. 1, pp. 549-578). Oxford: Oxford University Press.

Sonuga-Barke, E. J. S., Auerbach, J., Campbell, S. B., Daley, D., & Thompson, M. (2005). Varieties of preschool hyperactivity: multiple pathways from risk to disorder. *Developmental Science, 8*(2), 141-150.

Sonuga-Barke, E. J. S., Bitsakou, P., & Thompson, M. (2010). Beyond the Dual Pathway Model: Evidence for the Dissociation of Timing, Inhibitory, and Delay-Related Impairments in Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(4), 345-355.

Sonuga-Barke, E. J. S., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., et al. (2013). Nonpharmacological Interventions for ADHD: Systematic Review and Meta-Analyses of Randomized Controlled Trials of Dietary and Psychological Treatments. *American Journal of Psychiatry, 170*(3), 275-289.

Sonuga-Barke, E. J. S., Dalen, L., & Remington, B. (2003). Do executive deficits and delay aversion make independent contributions to preschool attention-deficit/hyperactivity disorder symptoms? *Journal of the American Academy of Child and Adolescent Psychiatry, 42*(11), 1335-1342.

Sonuga-Barke, E. J. S., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: A randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(4), 402-408.

Sonuga-Barke, E. J. S., Daley, D., Thomson, M., & Swanson, J. M. (2003). Preschool AD/HD: exploring uncertainties in diagnostic validity and utility, and treatment efficacy and safety. *Expert Review in Neurotherapeutics, 3*, 465-476.

Sonuga-Barke, E. J. S., De Houwer, J., De Ruiter, K., Ajzenstzen, M., & Holland, S. (2004). AD/HD and the capture of attention by briefly exposed delay-related cues: evidence from a conditioning paradigm. *Journal of Child Psychology and Psychiatry, 45*(2), 274-283.

Sonuga-Barke, E. J. S., & Halperin, J. M. (2010). Developmental phenotypes and causal pathways in attention deficit/hyperactivity disorder: potential targets for early intervention? *Journal of Child Psychology and Psychiatry, 51*(4), 368-389.

Sonuga-Barke, E. J. S., Houlberg, K., & Hall, M. (1994). When is "impulsiveness" not impulsive? The case of hyperactive children's cognitive style. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35*(7), 1247-1253.

Sonuga-Barke, E. J. S., Sergeant, J. A., Nigg, J., & Willcutt, E. (2008). Executive dysfunction and delay aversion in attention deficit hyperactivity disorder: Nosologic and diagnostic implications. *Child and Adolescent Psychiatric Clinics of North America, 17*(2), 367-384.

Sonuga-Barke, E. J. S., & Taylor, E. (1992). The effect of delay on hyperactive and non-hyperactive children's response times: a research note. *Journal of Child Psychology and Psychiatry, 33*(6), 1091-1096.

Sonuga-Barke, E. J. S., Taylor, E., & Heptinstall, E. (1992). Hyperactivity and delay aversion II: the effect of self versus externally imposed stimulus-presentation periods on memory. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(2), 399-409.

Sonuga-Barke, E. J. S., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion I: The effect of delay on choice. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33*(2), 387-398.

Sonuga-Barke, E. J. S., Thompson, M., Abikoff, H., Klein, R., & Brotman, L. M. (2006). Nonpharmacological interventions for preschoolers with ADHD - The case for specialized parent training. *Infants and Young Children, 19*(2), 142-153.

Sonuga-Barke, E. J. S., Thompson, M., Stevenson, J., & Viney, D. (1997). Patterns of behaviour problems among pre-school children. *Psychological Medicine, 27*(4), 909-918.

Sonuga-Barke, E. J. S., Wiersema, J. R., van der Meere, J. J., & Roeyers, H. (2010). Context-dependent Dynamic Processes in Attention Deficit/Hyperactivity Disorder: Differentiating Common and Unique Effects of State Regulation Deficits and Delay Aversion. *Neuropsychology Review, 20*(1), 86-102.

Strayhorn, J. M. (2002a). Self-Control: Theory and research. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1), 7-16.

Strayhorn, J. M. (2002b). Self-control: Toward systematic training programs. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1), 17-27.

Stroehle, A., Stoy, M., Wrase, J., Schwarzer, S., Schlagenhauf, F., Huss, M., et al. (2008). Reward anticipation and outcomes in adult males with attention-deficit/hyperactivity disorder. *NeuroImage, 39*(3), 966-972.

Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale, NJ: L. Erlbaum.

Supplee, L. H., Shaw, D. S., Hailstones, K., & Hartman, K. (2004). Family and child influences on early academic and emotion regulatory behaviors. *Journal of School Psychology, 42*(3), 221-242.

Swanson, J. M., Sergeant, J., Taylor, E., Sonuga-Barke, E. J. S., Jensen, P. S., & Cantwell, D. P. (1998). Attention-deficit hyperactivity disorder and hyperkinetic disorder. *The Lancet, 351*, 429-433.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Taylor, E., Dopfner, M., Sergeant, J., Asherson, P., Banaschewski, T., Buitelaar, J., et al. (2004). European clinical guidelines for hyperkinetic disorder - first upgrade. *European Child & Adolescent Psychiatry, 13*, I7-I30.

Taylor, E., & Sonuga-Barke, E. J. S. (2008). Disorders of Attention and Activity. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. S. Stevenson, E. A. Taylor & A. Thapar (Eds.), *Rutter's Child and Adolescent Psychiatry (5th Ed.)*: Wiley-Blackwell Publishers.

Thompson, M. J. J., Laver-Bradbury, C., Ayres, M., Le Poidevin, E., Mead, S., Dodds, C., et al. (2009). A small-scale randomized controlled trial of the revised new forest parenting programme for preschoolers with attention deficit hyperactivity disorder. *European Child & Adolescent Psychiatry, 18*(10), 605-616.

Thorell, L. B. (2007). Do delay aversion and executive function deficits make distinct contributions to the functional impact of ADHD symptoms? A study of early academic skill deficits. *Journal of Child Psychology and Psychiatry, 48*(11), 1061-1070.

Thorell, L. B., Lindqvist, S., Nutley, S. B., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. [Article]. *Developmental Science, 12*(1), 106-113.

Thorell, L. B., & Wåhlstedt, C. (2006). Executive Functioning Deficits in Relation to Symptoms of ADHD and/or ODD in Preschool Children. *Infant and Child Development, 15*(5), 503-518.

Toplak, M. E., Dockstader, C., & Tannock, R. (2006). Temporal information processing in ADHD: Findings to date and new methods. *Journal of Neuroscience Methods, 151*(1), 15-29.

Tripp, G., & Alsop, B. (2001). Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). [Article]. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(5), 691-698.

Tsukayama, E., & Duckworth, A. L. (2010). Domain-specific temporal discounting and temptation. *Judgment and Decision Making, 5*(2), 72-82.

Ursache, A., Blair, C., & Raver, C. C. (2012). The Promotion of Self-Regulation as a Means of Enhancing School Readiness and Early Achievement in Children at Risk for School Failure. *Child Development Perspectives, 6*(2), 122-128.

van der Meere, J. (2005). State regulation and attention deficit hyperactivity disorder. In D. Gozal & D. L. Molfese (Eds.), *Attention deficit hyperactivity disorder: from genes to patients*. Totowa, NJ: Humana Press.

van Meel, C. S., Heslenfeld, D. J., Oosterlaan, J., Luman, M., & Sergeant, J. A. (2011). ERPs associated with monitoring and evaluation of monetary reward and punishment in children with ADHD. *Journal of Child Psychology and Psychiatry, 52*(9), 942-953.

Vaughan, B. S., & Kratochvil, C. J. (2006). Pharmacotherapy of ADHD in young children. *Psychiatry, 3*(8), 36-44.

Vitiello, B. M. D., Severe, J. B. M. S., Greenhill, L. L. M. D., Arnold, L. E. M. D., Abikoff, H. B. P. D., Bukstein, O. G. M. D., et al. (2001). Methylphenidate Dosage for Children With ADHD Over Time Under Controlled Conditions: Lessons From the MTA. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 188-196.

Wang, X. T., & Dvorak, R. D. (2010). Sweet Future: Fluctuating Blood Glucose Levels Affect Future Discounting. *Psychological Science, 21*(2), 183-188.

Webster-Stratton, C. H., & Reid, M. J. (2010). The Incredible Years Program for Children from Infancy to Pre-adolescence: Prevention and Treatment of Behavior Problems. In R. C. Murrihy, A. D. Kidman & T. H. Ollendick (Eds.), *Clinical Handbook of Assessing and Treating Conduct Problems in Youth* (pp. 117-138).

Wiersema, R., van der Meere, J., Roeyers, H., Van Coster, R., & Baeyens, D. (2006). Event rate and event-related potentials in ADHD. *Journal of Child Psychology and Psychiatry, 47*(6), 560-567.

Wilbertz, G., Trueg, A., Sonuga-Barke, E. J. S., Blechert, J., Philipsen, A., & Tebartz van Elst, L. (2013). Neural and psychophysiological markers of delay aversion in attention-deficit hyperactivity disorder. *Journal of abnormal psychology, 122*(2), 566-572.

Wilcox, R. R. (2001). *Fundamentals of modern statistical methods : substantially improving power and accuracy*. New York: Springer.

Wilhelm, C. J., & Mitchell, S. H. (2009). Strain differences in delay discounting using inbred rats. *Genes Brain and Behavior, 8*(4), 426-434.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the Executive Function Theory of Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. *Biological Psychiatry, 57*(11), 1336-1346.

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

Willoughby, M. T., Kupersmidt, J. B., & Voegler-Lee, M. E. (2012). Is preschool executive function causally related to academic achievement? *Child Neuropsychology, 18*(1), 79-91.

Wilson, V. B., Mitchell, S. H., Musser, E. D., Schmitt, C. F., & Nigg, J. T. (2011). Delay discounting of reward in ADHD: application in young children. *Journal of Child Psychology and Psychiatry, 52*(3), 256-264.

Yi, R., Mitchell, S. H., & Bickel, W. K. (2010). Delay discounting and substance abuse dependence. In G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: the behaivoural and neurological science of discounting*. Washington, D.C.: American Psychological Association.

Zelazo, P. D., & Lyons, K. E. (2012). The Potential Benefits of Mindfulness Training in Early Childhood: A Developmental Social Cognitive Neuroscience Perspective. *Child Development Perspectives, 6*(2), 154-160.

Zentall, S. S., & Zentall, T. R. (1976). Activity and task-performance of hyperactive-children as a function of environmental stimulation. *Journal of Consulting and Clinical Psychology, 44*(5), 693-697.