Exploring the behaviour and wellbeing of UK schoolchildren using multi-informant reports

Dr Fiona Louise Wright BSc (Hons) PhD

Submitted in accordance with the requirements for the degree of
Doctor of Clinical Psychology (D. Clin. Psychol.)

The University of Leeds

School of Medicine

Academic Unit of Psychiatry and Behavioural Sciences

September, 2018

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

This work has been possible, only because of the collective skill, knowledge and experience of all involved. I am very grateful to all the children, as well as their parents and teachers, who have participated in this research. I would also like to thank the staff from the schools that took part, for being so enthusiastic and supportive of research.

I would like to thank my supervisors, Louise Dye and Clare Lawton, for providing me the chance to join their team and open up the opportunities for this thesis. I would also like to express my gratitude to the DClinPsyc course staff for their support, feedback and encouragement.

A massive thank you to my parents, Trish and Dave, for inspiring me to keep learning and studying. Your love and guidance is with me in whatever I chose to pursue. My sincere thanks also goes to my sister, Claire, whose respect continually reassures me of my potential.

Thank you to my fellow doctoral students and my friends, for all the fun we have had over the last three years and for always offering that glimmer of hope for post-thesis normalcy. You represent some of the most genuine, compassionate and thoughtful people I have ever had the pleasure of being with. Your honesty and humour have made this experience unforgettable.

Nobody has been more significant to me in the pursuit of this achievement than my wonderful husband, Will. Thank you (again) for the enduring love and encouragement you have shown me throughout this process. You have been the reminder to me of what is truly important in life.

"I almost wish I hadn't gone down that rabbit-hole – and yet – and yet – it's rather curious, you know, this sort of life!" – Alice in Wonderland.

Abstract

It has been suggested that polyphenols are linked to behaviour and wellbeing. A randomised chronic 10-week, parallel group, double blind, placebo controlled study investigated the effects of daily consumption of a flavanone-rich orange juice drink compared with a placebo on behaviour and wellbeing in 8 to 10 year old healthy schoolchildren (n=32). As the intervention study aimed to use data collected, from multiple informants using measures of behaviour and wellbeing, this thesis also aimed to establish the value of self-report measures in the accurate reflection of a child's behavioural and emotional difficulties, using the Strengths and Difficulties Questionnaire (SDQ). The tool development work explored the psychometric properties of SDQ data collected from different informants; parents, teachers and the children themselves, in a sample of healthy young schoolchildren aged 8 to 10 years old (n=126). The study found that the internal reliability, inter-rater agreement and concurrent validity of the SDQ, when completed by the adult informants (parents and teachers) were satisfactory and consistent with previous research. The structural validity of the teacher informant data supported the original five-factor structure of the SDQ as proposed by Goodman (1994). While, the parent informant data generated a four-factor solution and the children's self-reported data produced a three-factor solution, which appeared to reflect young children's categorisation of behaviours and traits. Overall, the factor structure generated provides valid information about behaviour and wellbeing in children as young as 8 years old.

The randomised chronic 10-week, parallel group, double blind, placebo controlled study, then used the tool development to investigate the effects of daily consumption of a flavanone-rich orange juice drink compared with a placebo on behaviour and wellbeing (using the SDQ, RCADS and SNAP-IV) in 8 to 10 year old healthy schoolchildren (n=32). This exploratory study did not find any effect of the flavanone-rich dietary intervention on any measures of behaviour and wellbeing, as reported by any of the informants. Taken together, the research presented in this thesis highlights the importance of assessing the validity and reliability of measures used to assess the behaviour and wellbeing of young children in clinical practice and research. It also provides an indication of how these measures might be further developed for future work in the field.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADHD	Attention Deficit and Hyperactivity Disorder
ANOVA	Analysis of Variance
ASD	Autism Spectrum Disorder
BASC	Behaviour Assessment System for Children
BDNF	Brain-Derived Neurotrophic Factor
CAM	Complementary and Alternative Medicine
CAMHS	Child and Adolescent Mental Health Services
CFA	Confirmatory Factor Analysis
CNS	Central Nervous System
CORC	Child Outcomes Research Consortium
CY-IAPT	Children and Young People's Improving Access to Psychological Therapies
d	Cohen's effect size
DSM	Diagnostic and Statistical Manual
EAL	English as an Additional Language
EFA	Exploratory Factor Analysis
FSM	Free School Meal
GAD	Generalised Anxiety Disorder
HADS	Hospital Anxiety and Depression Scale
HPA	Hypothalamic-Pituitary-Adrenal
IBS	Irritable Bowel Syndrome
IQ	Intelligence Quotient
MASC	Multidimensional Anxiety Scale for Children
MDD	Major Depressive Disorder
MIA	Maternal Immune Activation
OCD	Obsessive Compulsive Disorder
ODD	Oppositional Defiance Disorder
PCA	Principal Component Analysis
PD	Panic Disorder
PROMs	Patient Recorded Outcome Measures
QoL	Quality of Life
RCADS	Revised Child Anxiety and Depression Scale
RCT	Randomised Control Trial
SA	Separation Anxiety
SCARED	Screen for Child Anxiety Related Emotional Disorders
SDQ	Strengths and Difficulties Questionnaire
S.E	Standard Error of the Mean
SEN	Special Educational Needs
SNAP-IV	Swanson, Nolan and Pelham ADHD Questionnaire (version IV)
SoPREC	School of Psychology Research Ethics Committee
SP	Social Phobia
WISC	Wechsler Intelligence Scale for Children
YAM-5	Youth Anxiety Measure for DSM-5

Chapter One: An Introduction to the Effects of Diet and Dietary Interventions on Behaviour and Wellbeing

The evidence base supporting a link between diet and our functioning is well-established (Bellisle et al., 1998). The associations between diet and health have promoted the exploration of nutrition as a possible therapeutic means to enhance cognitive function and mood. The benefits of this type of approach include the cost-effectiveness and non-invasive nature of dietary intervention, and importantly, the fewer perceived or real side effects, relative to many pharmaceutical treatments. This thesis examines the potential for diet to influence behaviour and wellbeing, while considering the development of appropriate measures for this research.

This thesis formed part of a randomised controlled trial (RCT) designed to assess the effects of polyphenol rich orange juice on (1) cognition, (2) behaviour and wellbeing and (3) gut microbiota in healthy young children. The focus of this thesis is behaviour and wellbeing and the other outcomes (cognition and gut microbiota) are reported elsewhere.

1.1 Dietary Interventions and Cognition in Children

A number of studies have found that dietary interventions are associated with measures of cognition in both experimental and epidemiological settings (Reichelt, Westbrook, & Morris, 2017). Of relevance to the work presented in this thesis, dietary interventions have been found to improve *children's* performance on cognitive tasks. For example, Muthayya et al., (2007) found that an iso-caloric dietary intervention attenuated decline on immediate and delayed recognition tasks in Indian children of low socioeconomic status (LSES) compared to controls. A further *acute* intervention study found that British children aged 8 to 10 years given a single dose of a blueberry drink showed improvements on tests of memory and attention (Whyte & Williams, 2015). Other research comes from Baumgartner et al., (2012) who found beneficial *chronic* effects (8.5 months) of nutrient supplementation (omega-3 fatty acid and iron) on cognition in South African children aged 6 to 11 years old.

However, it is important to note that interventions which correct a nutritional deficiency are more likely to show beneficial effect. For example, a recent meta-analysis demonstrated that correcting iron deficiency results in improved cognitive function in anaemic children aged 6 years and over but not in iron depleted children or young women (Falkingham et al., 2010). Conversely, dietary intervention studies in healthy well-nourished children may also report cognitive benefits, although this may be related to the prevention of impairment during

cognitive demand rather than enhancement of performance as described above in children of poorer nutritional status. For example, studies have demonstrated that provision of breakfast prevents a decline in performance over the course of the morning in healthy children and adolescents compared to no breakfast (Hoyland, Dye, & Lawton, 2009) and that regular consumption of breakfast is associated with better academic outcomes (Adolphus, Lawton, & Dye, 2013; Adolphus, Lawton, & Dye, 2015). Nevertheless, Hoyland, Lawton, and Dye (2008) reported that the beneficial cognitive effects of breakfast consumption are most clearly demonstrated in children with nutritional deficiencies.

1.2 Dietary Interventions and Behaviour in Children

Laboratory based experimental studies have generally found that eating breakfast has a positive effect on children's cognitive performance, particularly in the domains of memory and attention (Cooper, Bandelow, & Nevill, 2011; Pivik, Tennal, Chapman, & Gu, 2012; Wesnes, Pincock, Richardson, Helm, & Hails, 2003; Wesnes, Pincock, & Scholey, 2012; Widenhorn-Müller, Hille, Klenk, & Weiland, 2008). In contrast, the study of the effects of dietary interventions, such as breakfast, on children's behaviour is comparatively rare (a recent meta-analysis identified only 12 studies: Adolphus et al., 2013). The paucity of research may be due to the limited measures available to assess children's behaviour and the need to develop standardized, validated, and comparable coding systems to measure behaviour has been documented . The majority of studies rate "on-task" and "off-task" behaviours usually in the classroom, which is a non-validated and unstandardized coding method. One of the criticisms of this method is its subjective nature and reliance on interpretation of behaviour. A systematic review by Adolphus et al. (2013) found that nineteen studies have employed behavioural measures to examine the effects of breakfast on behaviour at school between the years 1950 and 2013. Twelve of these used questionnaires and rating scales to measure the children's behaviour. Measures included: the Strength and Difficulties Questionnaire (SDQ; see section 2.5.1), Social Skills Rating System (SSRS), Child Behavior Checklist (CBCL) Conners Teacher Rating Scale (CTRS), and The Attention Deficit Disorder—Hyperactivity Comprehensive Teacher's Rating Scale (ACTeRS). Studies with low SES and undernourished young children, aged 8 to 10 years old, found beneficial effects on hyperactivity, using the CTRS (Murphy et al., 1998; Richter, Rose, & Griesel, 1997). However, no changes in SDQ outcomes were found following a one year intervention of universal free healthy breakfast provision (Mhurchu et al., 2013; Murphy et al., 2011).

1.3 Dietary Intervention and Psychological Functioning in Children

It is important to acknowledge that cognitive, behavioural, psychological and academic outcomes are not independent. For example, changes in cognitive performance, such as increases in attention, may be reflected by an increase in on-task behaviour during lessons and ultimately performance on academic outcome measures such as exams/in class tests. Similarly, changes in psychological wellbeing may also impact cognitive performance and behaviour in a cumulative manner.

There is correlational evidence linking diet and psychological functioning (Beezhold, Radnitz, Rinne, & DiMatteo, 2015; Crichton, Bryan, Hodgson, & Murphy, 2013). Numerous studies have shown that eating more fruit and vegetables is associated with elevated mood in adolescents, adults and older adults (Akbaraly, Sabia, Shipley, Batty, & Kivimaki, 2013; Conner, Brookie, Richardson, & Polak, 2015; Jacka et al., 2011; Le Port et al., 2012; José A Piqueras, Kuhne, Vera-Villarroel, Van Straten, & Cuijpers, 2011; Rienks, Dobson, & Mishra, 2013; Skarupski, Tangney, Li, Evans, & Morris, 2013; White, Horwath, & Conner, 2013). Furthermore, epidemiological studies have found that some diets (for example, those high in fruit, vegetables, fish, olive oils, nuts and legumes – such as the Mediterranean diet) can protect against depression (Lai et al., 2013; Psaltopoulou et al., 2013; Rienks et al., 2013; Ruusunen et al., 2014; Sánchez-Villegas et al., 2013; Sanhueza, Ryan, & Foxcroft, 2013; Skarupski et al., 2013). Correspondingly, depression and anxiety are associated with high-energy (fast-food) dietary patterns in adults (Hirth, Rahman, & Berenson, 2011; Sánchez-Villegas et al., 2009; Sánchez-Villegas et al., 2011).

There have also been a number of randomised control trials (RCTs) demonstrating the impact of whole-diet interventions on mood, for example depression and anxiety (Einvik et al., 2010; García-Toro et al., 2012; Wardle et al., 2000). A systematic review (Opie, O'Neil, Itsiopoulos, & Jacka, 2015) found that just less than 50% of RCTs investigating this link found significant improvements in depression (with Cohen's d or effect size between 0.19 and 2.02) and 20% of RCTs found significant improvements in anxiety. Furthermore, there are a number of studies investigating specific diet interventions (for example; cherries; Garrido et al., 2012; fish; Hansen et al., 2014; tomato juice; Hirose et al., 2015; and yoghurt; Jaatinen et al., 2014) and their positive impact on physical and psychological health in adults.

Specific micronutrient interventions have also been explored. For example, while there is pre-clinical evidence demonstrating that dietary interventions (for example, fish oil) reduce anxiety and improve cognition in non-human primates (Vinot et al., 2011), specific omega-3

supplementation has been found to improve symptoms of depression and bipolar disorder in adults (Nemets, Stahl, & Belmaker, 2002; Peet & Horrobin, 2002; Puri, Counsell, Richardson, & Horrobin, 2002; Stoll et al., 1999). The Oxford-Durham study, explored the effects of a 3-month omega-3 intervention on teacher-rated measures of inattention and emotional liability in schoolchildren and found reductions on CRTS-L (Connors Teaching Rating Scales – Long version) sub-scales following active treatment. However this research has been considerably criticised for it's "inadequate research methods" (Goldacre, 2006).

There are relatively few studies exploring the association in young children and adolescents. The majority of which are correlational by design. Khalid, Williams, and Reynolds (2016) conducted a systematic review and found twenty studies supporting an overall association between dietary patterns and mental health. It was also acknowledged that important confounding factors, such as socioeconomic status (SES) were rarely controlled for and the effect sizes were often small. Correspondingly, in another recent review, O'Neil et al. (2014) found evidence for a link between unhealthy diet and worsening mental health in children and adolescents. The 'western' dietary pattern has been implicated in poorer behaviour and wellbeing outcomes for adolescents (Oddy et al., 2009).

1.4 Polyphenols and Psychological Functioning

Polyphenols (technically polyhydroxyphenols) are a group of compounds found naturally in plants and are therefore, a basic component of our diet (Bravo, 1998; Trebatická & Ďuračková, 2015). They are thought to protect plants from ultra-violet radiation, disease and physical damage, and they may also activate a number of intracellular processes that preserve neurons, in humans (Gomez-Pinilla & Nguyen, 2012).

Chemically, polyphenols are comprised of multiple phenolic groups based on the number of phenol groups and the attached substances. Flavonoids (a 2-phenyl-1,4-benzopyrone structure) make up the most common group (see Figure 1). Flavonoid consumption has been linked to vascular and cognitive health (Macready et al., 2009; Miller & Shukitt-Hale, 2012). A number of mechanisms are being explored to understand these beneficial effects of flavonoids on functioning and health. Some of which include increases in cerebral blood flow (Lamport, Pal, et al., 2016; Vauzour, Vafeiadou, Rodriguez-Mateos, Rendeiro, & Spencer, 2008) and improving neural signalling pathways, such Brain-Derived Neurotrophic Factor (BDNF: Rendeiro et al., 2013).

Polyphenol absorption in the small intestine is low (Tuohy, Conterno, Gasperotti, & Viola, 2012) with as much as 90% of these compounds persisting into the colon. There, they are extensively metabolized (Selma, Espin, & Tomas-Barberan, 2009), by gut microbiota (including Bifidobacterium sp., Lactobacillus sp., Bacteroidete sp., Eubacterium sp. and

E.coli) resulting in smaller metabolites such as phenolic acids, some of which can then be absorbed across the intestinal mucosa and influence the gut-brain axis (Zhang et al., 2015). Although the exact mechanism is unclear, pre-clinical evidence has suggested that polyphenol intake can reduce anxiety-like behaviour in rodents (Allam et al., 2013; Harsha & Anilakumar, 2013; Scalbert & Williamson, 2000).

Epidemiological data shows that lifetime consumption of fruit and vegetables (and therefore higher flavonoid consumption) predicts a lower incidence of depression in later life (Bondonno et al., 2014; Bouayed, 2010; Chang et al., 2016; Mihrshahi, Dobson, & Mishra, 2015; Pase et al., 2013). Of particular interest has been cocoa (the main constituent of dark chocolate), which contains a complex mixture of polyphenols, especially flavanols (a subclass of the commonly ingested class of polyphenols called flavonoids; see Figure 1). Chocolate has been reported to ameliorate pre-existing negative mood states (Scholey & Owen, 2013). Sathyapalan, Beckett, Rigby, Mellor, and Atkin (2010) found that a cocoa intervention reduced self-reported anxiety and depression (measured using the Hospital Anxiety Depression Scale; HADS) in a people with chronic fatigue syndrome. Cocoa polyphenols have also been shown to improve self-reported mood (assessed using the Bond and Lader Visual Analogue Scale) and cognition in healthy adults (Pase et al., 2013).

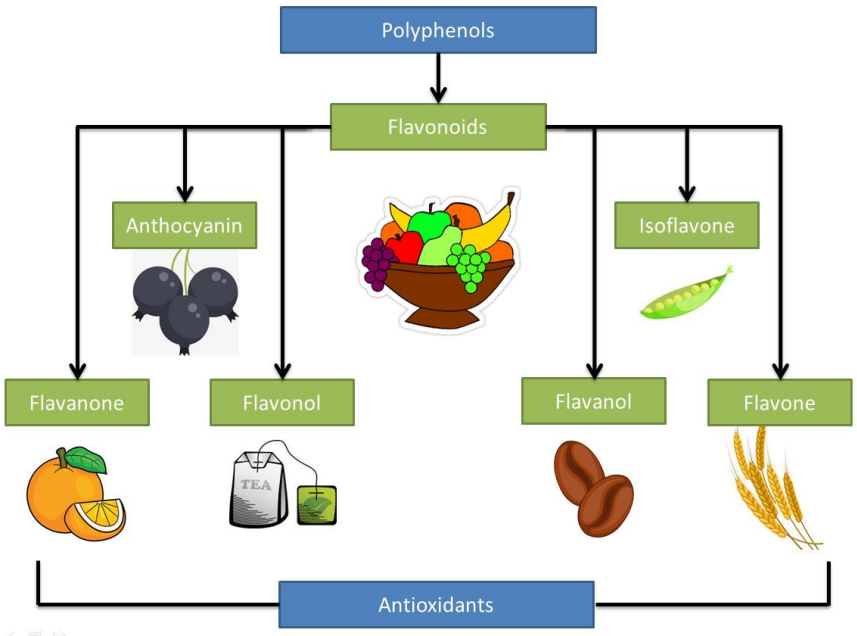


Figure 1: Understanding flavonoid polyphenols (adapted from: Szliszka & Krol, 2011)

There are high levels of flavonoids present in both green and black tea, which have also been found to reduce symptoms of depression and anxiety (Gardner, Ruxton, & Leeds, 2007). It is thought that tea polyphenols modulate levels of the monoamine oxidase enzyme,

which in turn increases monoamine concentrations (Mazzio, Harris, & Soliman, 1998). This is consistent with evidence that curcumin (found in turmeric) exerts its antidepressant effect, in mice, via modulation of the expression of brain serotonergic and dopaminergic neurotransmission (Kulkarni, Bhutani, & Bishnoi, 2008). Similarly, a number of human studies have found an impact of soy-derived isoflavones on mood and cognition in menopausal women (Chedraui, San Miguel, & Schwager, 2011; Hirose et al., 2016; Kritz-Silverstein, Von Mühlen, Barrett-Connor, & Bressel, 2003).

Recent studies have found acute benefits of increased positive affect following the consumption of a flavonoid-rich wild blueberry drink in healthy young children (aged 7 to 10 years old; Whyte, Schafer, & Williams, 2016). However, to date, few studies have examined the effects of a polyphenol-rich orange juice intervention on behaviour and wellbeing. One study in healthy older adults over 8 weeks, found no significant changes in anxiety or depression symptoms using the HADS (Kean et al., 2015). Similarly, Alharbi et al. (2016) reported an acute (6 hour) improvement in alertness and cognitive performance in healthy middle-aged male adults following flavonoid-rich orange juice. Yet, there were no significant effects on mood ratings.

1.4.1 Possible Mechanisms of Action

One possible explanation for the association between the consumption of flavonoids and decreased risk of depression and mood disorders, is the beneficial effects of flavonoids on executive functioning. Impaired cognitive functioning has been linked to the maintenance of depression and low mood (Gohier et al., 2009; Marazziti, Consoli, Picchetti, Carlini, & Faravelli, 2010; McDermott & Ebmeier, 2009; Naismith et al., 2003), while there are a number of studies which support the association between flavonoids and improved cognitive and physiological health (Bell, Lamport, Butler, & Williams, 2015).

It has been suggested that improvements observed in cognitive functioning, following high-flavanone drinks (in healthy adults), may be associated with enhanced blood flow to the brain (Lamport, Pal, et al., 2016; Vauzour et al., 2008). Additionally, it is thought that polyphenols act via a reduction in oxidative stress, which is associated with many neurological and psychiatric disorders (McMartin, Jacka, & Colman, 2013; Zhang et al., 2015). Oxidative stress, an inflammatory response, is an over production of free radicals leading to oxidative damage to lipids and proteins, resulting in the dysfunction of cells and organs. It is thought that, polyphenols may exert their neuro-protective effects by modulating specific signalling pathways involved in cognitive processes and synaptic plasticity (Gomez-Pinilla & Nguyen, 2012). Polyphenol intake has been found to increase levels of BDNF (Brain-Derived Neurotrophic Factor), which is associated with synaptic plasticity (Zuccato & Cattaneo, 2009). Interestingly, levels of BDNF are often found to be

altered in depression and schizophrenia (Angelucci, Brene, & Mathe, 2005), which is consistent with evidence suggesting that Resveratrol (a polyphenol derivative found in grapes) exerts its antidepressant effect via BDNF (Hurley, Akinfiresoye, Kalejaiye, & Tizabi, 2014). While, cocoa flavonoids have been suggested to modulate psychological functioning via the normalisation of gut microbial activities (Martin & Appel, 2009).

Another possible mechanism of action for effects on mood is the impact of polyphenols on monoamine oxidase activity. Watson et al. (2015), recently, provided a description of the inhibition of monoamine oxidase-A and monoamine oxidase-B, alongside cognitive benefits, using an acute polyphenol-rich blackcurrant supplement. Interestingly, monoamine oxidase-B has been used for the treatment of depressive disorders (Youdim & Bakhle, 2006), which may support the link between mood, cognition and behaviour. Similarly, the discovery of flavonoids as benzodiazepine binding site ligands promoted research into the role of flavonoid's anxiolytic-like effects via the GABA pathways (Wasowski & Marder, 2012).

1.5 Dietary Interventions and Clinical Psychology

Clinical Psychologists receive no formal training on diet/nutrition and therefore, are ill-equipped to discuss diet and secondary life-style factors associated with mental health with clients. However, acknowledging the multifactorial basis of psychopathology can be an effective and relatively simple way to promote mental health and recovery from mental illness that can be easily integrated into healthcare (Lachance & Ramsey, 2015). The potential implications of findings that a dietary intervention which increases flavonoid consumption can attenuate mood disorders would be significant. Furthermore, a number of psychological difficulties tend to develop during adolescence or early adulthood, which often extend into adulthood or re-emerge as a relapse in later life. Therefore, intervention during critical periods of development in young children could potentially have enormous implications for the development of dietary habits (such as polyphenol intake: Birch & Fisher, 1998) and the incidence of mental health difficulties. It is important to explore the links between diet and mental health and translate any findings into practical, cost-effective and acceptable interventions for use in the NHS.

Some areas of research have contributed to establishing a strong link between diet and psychological symptomology (Louis, 2012). Autism Spectrum Disorders (ASD), are pervasive neurodevelopmental disorders that are characterised by difficulties in social interaction and communication. Additionally, gastrointestinal symptoms are often reported, which may suggests a role of diet (Buie et al., 2010) and abnormal gut microbiota composition in ASD (Adams, Johansen, Powell, Quig, & Rubin, 2011; Mulle, Sharp, &

Cubells, 2013). Furthermore, many children with ASD practice Complementary and Alternative Medicine (CAM) in the form of dietary interventions (Levy & Hyman, 2003), for example pre- and pro-biotic treatment (de Theije et al., 2011).

1.5.1 An Example of Dietary Intervention and Clinical Psychology: Attention Deficit Hyperactivity Disorder (ADHD)

Historically, numerous studies have alleged and/or demonstrated an effect of nutritional elements of psychopathological symptomology. For example, it has been hypothesised that sugar and/or food additives are associated with hyperactivity (Crook, 1974; Feingold, 1975), specifically the condition known as "attention-deficit/hyperactivity disorder" (ADHD; Wender & Solanto, 1991). Although the current research, using a community sample of young children, is not directly exploring the impact of dietary interventions as an ADHD treatment, it is relevant to consider the effects of these interventions on reports of behaviours associated with ADHD (such as inattention and hyperactivity) in a community sample.

There is evidence to suggest that dietary interventions (for example, elimination, food colour exclusion and macronutrient supplementation) can improve psychological functioning in study participants meeting the criteria for ADHD (see reviews: Heilskov Rytter et al., 2015; J. Rucklidge, Taylor, & Whitehead, 2011; Sonuga-Barke et al., 2013). Additionally, it has been suggested that there is a significant role of dietary deficiency in ADHD, specifically an imbalance of essential/free fatty acid (omega-3 and -6; Bloch & Qawasmi, 2011). Similarly, it has been hypothesised that high-calorie 'junk food' diets are associated with behavioural problems (such as aggression; Benton et al., 1997). However, the experimental evidence does not support a causal effect. There are, therefore, various possible dietary factors suggested but little evidence. Importantly, it has been suggested that findings of dietary interventions should take into consideration the possible correction of poor nutritional state (Bellisle, 2004).

1.5.1.1 Polyphenols and ADHD

Interestingly, consumption of polyphenol rich food is associated with low incidence of ADHD (Trebatická et al., 2006), and might therefore be beneficial for kids who are hyperactive but do not meet the criteria for a diagnosis. Several polyphenolic compounds have been evaluated for the treatment of ADHD behaviours (Rucklidge, Johnstone, & Kaplan, 2009), specifically Pycnogenol® a standardised extract of French maritime pine bark (Pinus pinaster). Trebatická et al. (2006) found one month of Pycnogenol® caused reductions in hyperactivity and improvements in attention and coordination in 6 to 14 year old children with ADHD.

Taken together, the currently available literature suggests that a polyphenol intervention has the potential to affect behaviour and wellbeing. The effects of polyphenol (flavonoid)-rich orange juice on psychological functioning has been recently examined in adults (Alharbi et al., 2016; Kean et al., 2015) but not in children. Given that a flavonoid rich blueberry drink has been found to improve positive affect (using the child version of the Positive and Negative Affect Scale; PANAS-C) in 7 to 10 year old children (Khalid et al., 2017), it is likely that a flavonoid-rich orange juice might yield similar benefits in this age group.

Chapter Two: An Introduction to Measuring Behaviour and Wellbeing

The main aim of this thesis was to explore the effects of chronic consumption of flavanone-rich orange juice compared with placebo (orange flavoured drink) on behaviour and wellbeing measures in 8 to 10 year old schoolchildren. However, prior to addressing this research question, appropriate methodology should be chosen. It is important that research incorporates methods to establish the reliability and validity of the tools used. A failure to acknowledge this may lead to difficulties in interpreting the results, and this may in-turn influence the clinical implications.

2.1 Measures of Behaviour and Wellbeing in Healthcare Services

Service user involvement in health care services has progressed beyond seeking assurances of satisfaction, and now aims to promote collaboration. The government's vision of this within the NHS has been set out in the "no decision about me, without me" documentation (Assessment, 2010; Elwyn et al., 2012; Health, 2010; Lansley, 2010). These new policies recommend the use of patient reported outcome measures (PROMs). These are typically self-completed questionnaires that collect service users' perspectives of their experience by gathering quantitative data about their symptoms, their functional status and their health related quality of life.

PROMS are tools that were originally developed for use in research, but have since been adopted by clinicians and services to improve care. The adoption of these tools was intended to inform clinical practice, feed into service development and facilitate local and national benchmarking. Although the majority of clinicians seek to incorporate the views of their service users into their care, a proportion of clinicians question the appropriate and effective use of PROMs. For example, the majority of these measures give equal weighing to all domains and don't capture the idiosyncratic values and priorities of their clients and concerns have been raised around how these tools might inform decision-making around resource allocation (Dunckley, Aspinal, Addington-Hall, Hughes, & Higginson, 2005). Nevertheless, in recent years, there has been a push from the government for healthcare providers to use PROMs to assess and compare their outcomes (Black, 2013).

It is important for clinicians to take into account the psychometric properties of the measures that are distributed to specific populations to ensure that the tools used are appropriate for and relevant to the clients.

2.2 Using Clinical Measures within Children's Services

The drive for using PROMs is apparent across all health services, including child and adolescent mental health services (CAMHS). The Child Outcomes Research Consortium (CORC) is the UK organisation that aims to collect and use these measures to improve mental health and wellbeing for children and young people (CORC, Accessed 2018). Similarly, NHS England in partnership with Health Education England introduced the Children and Young People's Improving Access to Psychological Therapies (CPY-IAPT) programme which works with a number of healthcare providers to promote collaboration with young people through regular feedback and outcome monitoring (CYP-IAPT, Accessed 2018; Wolpert, Fugard, Deighton, & Görzig, 2012). In 2014, these groups published a guide on the use of outcome measures and feedback in everyday clinical practice with young people (Law & Wolpert, 2014). Factors to consider in the selection of outcome measures include; how easily they embed into services, the timely feedback needed to inform practice and their use in cross-service comparisons (Wolpert et al., 2014; Wolpert, Ford, et al., 2012). This has generally led to the use of standardised rating scales that measure global constructs, which can be completed by various individuals across multiple time points (see review: Myers & Winters, 2002a, 2002b). As rating scales are easy to score and interpret, specialist training is not required and they are, therefore, an efficient and economical option.

Outcome orientated CAMHS policies now recommend the use of brief questionnaires, which can be completed outside of the therapeutic sessions, often prior to a first appointment, to help clinicians understand a family's difficulties. These typically ask the clients to respond to questions or statements using a numerical rating. For example, "I have felt OK about myself this week" scored using a 5-point scale ranging from "0 – not at all" to "5 – most of the time". The multiple questions (also known as items) on the measure often cover a number of theoretical dimensions (for example, (1) wellbeing, (2) symptomology, (3) life-functioning or (4) risk). The responses to items relating to these dimensions (also known as sub-scales) are designed to be averaged by the clinician to produce a mean score. This can, in turn, be used to indicate the level of psychological distress or symptomology (for example, from 'healthy' to 'severe'). Using these scores as cut-offs for classification (such as diagnosis) has several advantages and disadvantages in clinical and research settings (Law & Wolpert, 2014).. They can also be used in review, and are given out after

six months and at the end of treatment, to track progress or change following clinical intervention.

These tools are also being incorporated into normative settings, such as schools or community groups, as screening tools to assess and monitor emotional wellbeing, with the hope of identifying the emergence of symptomology in high-risk young people. However, there are concerns that the practitioners using these measures are not appropriately trained to evaluate the outcomes in the context of an understanding of their psychometric properties. Similarly, there is a risk that the broad nature of these measures can mean that feedback does not provide sufficient detail prior to classification/diagnosis and can lead to mistaken decision making (Wolpert et al., 2014).

2.3 Methodological Issues in the Measurement of Behaviour and Wellbeing in Children: Self-report vs. Adult Informants

The focus on client collaboration means that national and international policy is focusing on the importance of the voice of the child, of shared decision making for children accessing health services and of self-defined recovery (Dex & Hollingworth, 2012; Government, 2004; Health, 2010; UNICEF, 1989) However, the assessment of mental health in young children has traditionally relied exclusively on adult informants, such as parents and/or teachers. Furthermore, these measures (also known as informant-rating scales) are not without disadvantage. There is much evidence to highlight parent and teacher biases, which often results in poor agreement in the reporting of symptoms (Galloway & Newman, 2017). Natural differences in different informants' experiences of the child's behaviour may explain these discrepancies. Garrison and Earls (1985) found that the highest agreement occurs for concrete, observable symptoms, while low agreement is shown for subjective judgments, about mental health. Parents often report more behavioural problems, while teachers report problems associated with academia and compliance (Garrison & Earls, 1985; Van Roy, Groholt, Heyerdahl, & Clench-Aas, 2010). Therefore, exclusive dependence on the reports of adults may be problematic, as these informants may fail to recognize the symptoms of emotional distress in children. There is also a concern that while different informants offer unique views of the child, this can be associated with bias and misreporting (Rescorla et al., 2014).

Consistent with recent policy, it would seem common-sense that the preferable source of health status information would be the patient, provided that reliable and valid data can be obtained. In this context, self-report measures could be particularly important for the assessment of emotional disorders in which self-disclosure of internal states may have greater validity than the ratings of other informants. Interestingly, evidence indicates that

when self-report is missing or dropped from assessment, the sensitivity to emotional disorders is reduced (Goodman, Meltzer, & Bailey, 2003). Nevertheless, concerns around self-report rating scales relate predominately to the psychometric properties of the available measures. Traditionally, measures for young people were modifications of adult scales and so were not developmentally appropriate (Kwan & Rickwood, 2015). More recently, scales have been developed specifically for young people, but concerns remain around young people's competence/ability to report on their own feelings and behaviours.

Similar to the assessment of mental health, quality of life (QoL) measurement in children also poses methodological problems. The changes in children's ability to understand at different ages, the difficulty in separating the child's perceptions from the parents', and the variation of activities with age, raises the question of whether reliable and valid self-report data can be obtained from children themselves (le Coq, Boeke, Bezemer, Bruil, & van Eijk, 2000). From a developmental perspective, the use of traditional self-report questionnaires in children younger than 12 years old has been questioned, and discouraged in children younger than 8 years (Stone, Otten, Engels, Vermulst, & Janssens, 2010). Children under 6 years of age have been regarded as developmentally unable to serve as valid reporters of their own mental state (Luby, Belden, Sullivan, & Spitznagel, 2007). With younger populations (late childhood: 8 to 10 years) it is often felt that the children might (1) not understand the language used, (2) have limited knowledge of the concepts being measured and (3) have difficulties in accurately comparing themselves to others (Denham, 1998).

2.3.1 Reliability of Self-report Measures in Children

Due to limited linguistic, cognitive, and social-emotional abilities, the ***reliability*** of self-reported symptomology in children is substantially lower than that of their parents. Yet, evidence suggests that the reliability of the child reports increase with age. Hence, adolescents were found to be as reliable as their parents in many symptom areas (Schwab-Stone, Fallon, Briggs, & Crowther, 1994).

A study assessing the reliability of the psychiatric interview of children, used a test-retest reliability of 0.70 as a criterion for reliable reporting and concluded that children aged 10 and above could self-report (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985). Similarly, Silverman and Eisen (1992) examined the reporting of anxiety symptomology in younger children and found that children aged 12 to 17 had better test-retest reliability than did children aged 6 to 11 years old, except when reporting avoidant and overanxious symptoms. Furthermore, the studies consistently found that questions about observable behaviours elicited more reliable responses than questions about emotions (Fallon & Schwab-Stone, 1994). It is possible that the reliability of the reports given by children reflect daily events rather than more stable patterns across time. Moreover, consistent

information across time is not necessarily indicative of a valid measure, but is certainly a prerequisite for validity. These studies used test-retest reliability to assess the appropriateness of self-report measures in children. This method assesses the instruments themselves; the reliability with which the reporting of symptom clusters may provide stability in the diagnosis of a disorder, from one time point to another (approx. 1-3weeks).

In contrast, an alternative approach to test-retest reliability is to assess the correlations between informants in order to assess the informant reliability; the reliability with which the responses from younger children converge with those of the teacher and parent. This is based upon the assumption that the adult informant responses represents the "true" response (or that the "truth" lies between the child and the parent).

2.3.2 Validity of Self-report Measures in Children

There are two ways to assess the ***validity*** of self-report measures; one possibility is to distinguish between a clinical and non-clinical sample. Another approach is to assess the correlations between scores on one measure with scores on other related measures, which may provide evidence for concurrent validity of the self-report measure.

Achenbach, Mcconaughy, and Howell (1987) carried out a meta-analysis of 119 studies assessing the behaviour and emotional problems of children. They found that the average (weighted) correlations between the child/adolescent self-reports and those of other informants were low (parent, 0.25; teacher 0.20). Interestingly, these correlations were significantly higher for 6 to 11 year-olds than for adolescents (aged 12 to 19). They also found that the correlations were lower for emotional problems such as anxiety and withdrawal (termed over-controlled), in contrast to problems such as aggression and hyperactivity (termed under-controlled). More recently, Jaureguizar, Bernaras, and Garaigordobil (2017) also found that the consistency between self-reports and teacher reports of depression in 7 to 10 year-olds was low. These data suggest that teachers might under-report depressive symptomology. Additionally, the measure used to assess depression may not be valid for use in younger populations, or indicate that child and adolescent behaviour and emotional problems cannot be effectively captured by informant-only measures. It has been suggested that there is an informant gradient, with self-report information from children having poorer screening properties than information from parents, whilst the combination of child and parent information is suggested to provide the best screening properties (Kuhn et al., 2017).

2.4 Self-report Measures of Behaviour and Wellbeing in Children

A review of self-reported physical and emotional symptoms (including scales for anxiety and depression) in cancer patients found that most studies described reliability and validity in a cohort in which most children were older than 8 years of age. Of the eight studies that evaluated reliability within the younger age group (younger than 8 years-old), all failed to demonstrate that the measures were reliable or valid (Pinheiro et al., 2017; Reeve et al., 2017). Similarly, Cremeens, Eiser, and Blades (2006) reviewed self-report measures for children, specifically those aged between 3 and 8 years, and found that there was a variety of measures available, but with shortcomings associated with reliability and validity.

In relation to anxiety, there are a number of well-established self-report measures for assessing symptoms of anxiety disorders, which are used clinically. These include; the Multidimensional Anxiety Scale for Children (MASC; March, Parker, Sullivan, Stallings, & Conners, 1997), the Screen for Child Anxiety Related Emotional Disorders (SCARED; Muris, Merckelbach, Schmidt, & Mayer, 1998), or the revised child anxiety and depression scale (RCADS; Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000). Most recently, Muris et al. (2017) developed the Youth Anxiety Measure for DSM-5 (YAM-5), a new scale for assessing child and adolescent anxiety disorders according to DSM-5 criteria (Garcia-Lopez, Saez-Castillo, & Fuentes-Rodriguez, 2017). However, these measures are primarily assessed using populations of adolescents aged 12 to 18 years old. In contrast, the Behaviour Assessment System for Children (BASC) was assessed with exclusively older adolescent children (16 to 18 year olds; Weis & Smenner, 2007).

Some self-report measures for children have been used/developed for younger samples. For example, the Children's self-report questionnaire was reviewed in 7 to 12 year old children (Beitchman & Corradini, 1988). Furthermore, Chorpita, Moffitt, and Gray (2005) validated the RCADS with 7 to 18 year olds and found that the mean age of the sample had an effect on the reliability for some sub-scales (Social Phobia, Panic Disorder, Generalised Anxiety Disorder, and Major Mood Disorder), such that higher mean age was associated with greater reliability coefficients (Piqueras, Martín-Vivar, Sandin, San Luis, & Pineda).

Worryingly, it seems that self-report measures are primarily designed and validated for teenagers and yet in practice, are often used with younger populations, sometimes with only minor modifications. It may be that clinical interviews and psychometric assessment with younger children are valuable for a number of reasons establishing rapport, ascertaining cognitive functioning, observing mental status, motor functioning and verbal skills), rather than a focus on the validity of the outcome obtained on a measure. However, they should not be taken at face value for purposes of diagnosis, treatment selection or, importantly for this study, treatment/intervention evaluation.

2.5 Measures of Behaviour and Wellbeing in Children Used in this Thesis

This thesis uses three outcome measures: the Strengths and Difficulties Questionnaire (SDQ), the Revised Child Anxiety and Depression Scale (RCADS) and the Swanson, Nolan and Pelham – IV Questionnaire (SNAP-IV). The SDQ and RCADS were selected as they are used nationally in CAMHS across the UK and represent standardised tools with a strong evidence based for use in service evaluation and direct clinical work. Both of these measures also include a self-report version. The SNAP-IV was chosen as it represents a short, freely available measure for assessing specific behavioural symptomology and is also regularly used in clinical practice to screen for Attention Deficit Hyperactivity Disorder (ADHD)-specific symptomology.

This overview will start with the SDQ, since this is the main focus of the thesis, and go on to discuss the RCADS (as this has an established self-report version) and finally the SNAP-IV.

2.5.1 The Strengths and Difficulties Questionnaire (SDQ)

2.5.1.1 An Overview of the Strengths and Difficulties Questionnaire (SDQ)

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) is a brief 25-item scale covering positive and negative attributes in five domains: (1) Emotional Difficulties, (2) Conduct Difficulties, (3) Hyperactivity-inattention Difficulties, (4) Peer Problems and (5) Prosocial Behaviour. Items are scored on a three-point scale (0=not true, 1=somewhat true and 2=certainly true) and summed to produce domain scores (Appendix 1). There is also an 'impact supplement' which can be used to assesses the overall distress and impairment perceived by the parent informant.

The measure is available in parent and teacher informed versions for children aged 4 to 17 years, with a self-report also available for children aged 11 to 17 years. It has an extensive peer-reviewed evidence base detailing its psychometric properties across a number of nations including Australia, China, Greece, Israel, Italy, Netherlands, Norway, Sweden, and United Kingdom (Di Riso et al., 2010; Du, Kou, & Coghill, 2008; Giannakopoulos et al., 2009; Haynes, Gilmore, Shochet, Campbell, & Roberts, 2013; Liu et al., 2013; Lundh, Wangby-Lundh, & Bjarehed, 2008; Mansbach-Kleinfeld, Apter, Farbstein, Levine, & Poznizovsky, 2010; Van Roy, Veenstra, & Clench-Aas, 2008; Yao et al., 2009)

The SDQ is a good measure of psychopathological symptoms in children and adolescents and correlates (aka. convergent validity) with other broad measures of psychopathology such as the Rutter (1967) and Achenbach (1991) questionnaires. It has demonstrated internal

consistency and test-retest reliability (Goodman, 2001). The measure shows moderate inter-rater agreement across parents, teacher and self-report (11-16yr old; Goodman et al., 2003), even in community samples. The SDQ has several advantages over other screening instruments, including its brevity and its focus on competencies as well as difficulties (Goodman 1997).

2.5.1.2 Prosocial SDQ Items

The SDQ developed as a modification of the Rutter parent questionnaire (Rutter, 1967), by including items on children's strengths (Goodman, 1994). Goodman found that when using the Rutter parent and teacher questionnaires in an assessment battery in an epidemiological study, many of the parents and teachers found the focus of the Rutter items to be disconcerting (Goodman 1994). Although it is common that screening tools focus on undesirable symptomology, informants can often feel that they want to identify the child's strengths as well as weaknesses. The inclusion of the prosocial items, influenced by the prosocial behaviour questionnaire (Weir and Duveen, 1981), reflected a desire to make the questionnaire more appealing and increase response rate. The findings suggested that the prosocial scale was a separate construct (rather an opposite representation of difficulties). This was consistent with psychiatric ideas at the time which focused on positive and protective factors of personality (Hay, 1994). To avoid response bias (for example, in the form of acquiescence), the final questionnaire includes statements reflecting both positive and negative attributes, including some items which have been reverse worded.

The first studies (Goodman, 1997) found that SDQ parent and teacher- informant responses for children aged 4 to 16 years old were highly correlated with those from the Rutter Questionnaire. Additionally, the new questionnaire did not differ in its ability to discriminate between psychiatric and dental clinic attenders, and interestingly, mothers of low-risk children were twice as likely to prefer the SDQ over the Rutter Questionnaire (Goodman & Scott, 1999).

2.5.1.3 SDQ Self-report Version

The self-report version of the SDQ was subsequently developed by Goodman in 1998. The early studies found that the inter-rater correlations between self-report, and parent and teacher informant responses were comparable to other measures, and the self-report version was able to discriminate between community and clinical samples (Goodman, Meltzer, & Bailey, 1998).

The *informant*-rated version of the SDQ can be completed by either the parents/guardians or teachers of children and teenagers aged between 4 and 16. However, the *self*-report version of the SDQ was intended for self-completion by young people aged between 11 and 16 (Van

Roy et al., 2008; van Widenfelt, Goedhart, Treffers, & Goodman, 2003). The 25 items of the self-report SDQ cover the same attributes as the informant-rated SDQ. For the majority of the items, the key difference between the *informant*-rated and *self*-rated version of the SDQ is a grammatical change from third to first person.

Most research on the SDQ has focused on upper primary school aged children and adolescents attending secondary school (11 to 17 years old; Stone et al., 2010). The psychometric properties of the self-report SDQ in these older children (13 to 17 year old) have been found sufficient in community (Koskelainen, Sourander, & Vauras, 2001) and clinical settings (Klasen et al., 2000), but research conducted on lower primary school-aged (8 to 11 years old) children shows mixed findings. The self-report measure has been used in a younger community sample of 8- to 13-year old children and showed sufficient internal consistency for Emotional and Hyperactivity, in addition to Prosocial Behaviour, domains was found (Koskelainen, Sourander, & Kaljonen, 2000; Muris, Meesters, Eijkelenboom, & Vincken, 2004). Nevertheless, the psychometric qualities of the self-report version in younger children were found to be reduced relative to children of 11 years and above. It is therefore recommended that the self-report measure is used in combination with versions completed by other informants (Muris et al., 2004). The evidence supported satisfactory inter-rater agreement and convergent validity with established measures.

2.5.1.4 The Five Factor Structure of the SDQ

The structure of the SDQ has been studied widely. The Goodman (1997) principal component analysis (PCA) produced 15 factors, but the authors decided that a rotated solution with six factors made the most clinical sense. The original Rutter questionnaire included scales relating to difficulties with Conduct, Emotion and Hyperactivity, therefore the factors for the SDQ were named Hyperactivity/Inattention (Factor I), Prosocial behaviour (Factor II), Conduct problems/oppositionality (Factor III), Somatic/developmental (Factor IV), Internalization (Factor V), and Peer relationships (Factor VI). In the newer version of the questionnaire, 'Internalisation' has been renamed 'Emotional Difficulties" and 'Somatic/developmental' has been removed. These changes resulted in the current five-factor structure of the SDQ.

The remaining five factors were substantiated in a British population by Goodman (2001). This structure has also been reported for French, German, Greek, Finish, Swedish, and, US populations, for the informant version (d'Acremont & Linden, 2008; Dickey & Blumberg, 2004; Giannakopoulos et al., 2009; Koskelainen et al., 2001; Smedje, Broman, Hetta, & von Knorring, 1999; Woerner et al., 2002) and the self-report version (Capron, Thérond, & Duyme, 2007; Giannakopoulos et al., 2009; Goodman, 2001; Hawes & Dadds, 2004; Koskelainen et al., 2001; Lundh et al., 2008; Muris et al., 2004; Van Roy et al., 2008).

However, other studies have found that the SDQ does not have a good fit for the reputed sub-scales (Mellor & Stokes, 2007) and have found support for three- (Di Riso et al., 2010; Koskelainen et al., 2001; Percy, McCrystal, & Higgins, 2008; Ruchkin, Jones, Vermeiren, & Schwab-Stone, 2008) and four- factor models (Bull, Lee, Koh, & Poon, 2016; Muris et al., 2004). The lack of uni-dimensionality of the sub-scales means that two or more factors are contributing to each scale, in turn resulting in unpredictable scale behaviour (Mellor & Stokes, 2007).

It has been suggested that a three-factor structure, representing (1) Prosocial behaviours, (2) Internalising and (3) Externalising problems may be an invariant structure across translations, informants and rotation strategies (Dickey & Blumberg, 2004). The Externalising score refers to problems directed outwards (such as Conduct and Hyperactivity) whilst the Internalising score refers to problems directed inwards (such as Emotional Difficulties and Peer Problems). Nevertheless, in clinical use, the five factor structure is used to produce five-sub-scale scores (Goodman, Lamping, & Ploubidis, 2010; Goodman, 2001).

2.5.2 Revised Child Anxiety and Depression Scale (RCADS)

The RCADS (Chorpita et al., 2000; Muris, Meesters, & Schouten, 2002) is a 45-item questionnaire representing six sub-scales: (1) Generalised Anxiety Disorder, (2) Separation Anxiety Disorder, (3) Social Phobia, (4) Panic Disorder, (5) Obsessive Compulsive Disorder and (6) Major Depressive Disorder. It also has a composite score to capture 'Total Difficulty'. The RCADS-parent version has demonstrated acceptable psychometric properties, including internal consistency, convergent/divergent validity and discriminant validity (Chorpita et al., 2005; Chorpita et al., 2000) .

While, the psychometric properties of the shortened 25-item (RCADS-25) self-report version are comparable to those obtained with the full-length version in a sample of normal schoolchildren aged 8- to 15-years old (Muris et al., 2002). The self-report version has the same six scales and it also has two composite scores: (1) Total Anxiety and (2) Total Difficulty (Appendix 1). The measure has demonstrated good internal consistency and test-retest reliability, in addition to showing reasonable correlations with parent reports and other self-report questionnaires (such as the STAIC and FSSC-R; Muris et al., 2002). Within this thesis, the RCADS is used to check the validity of the SDQ by comparing reports of the same construct of wellbeing and/or behaviour. For example, adult informed SDQ sub-scale scores will be compared with adult informed RCADS sub-scale scores, and self-reported SDQ sub-scale scores will be compared with both adult informed and self-report RCADS sub-scale scores.

2.5.3 Swanson, Nolan and Pelham - IV Questionnaire (SNAP-IV)

The SNAP-IV is a behaviour rating scale for the diagnosis of ADHD based on the DSM-III (Diagnostic and Statistical Manual of Mental Disorders, version III). The use of SNAP-IV within this thesis was recommended by a CAMHS contact. The SNAP-IV is a free resource that compares favourably to the more resource intensive Connors Test (Kao & Thomas, 2010). The study reported within this thesis used a short, 18-item informant-report version of the SNAP-IV to assess symptoms of inattentiveness and hyperactivity/impulsiveness, whereby items are scored on a three-point scale (0=Not at all, 1=Just a little, 2=Quite a bit and 3=Very much). This is an adaptation of the 26-item MTA (Multimodal Treatment of ADHD), excluding items relating to oppositional defiance disorder (ODD; Swanson et al., 2001). The SNAP-IV has been criticised for a lack of published psychometric properties and a paucity of normative data (Collett, Ohan, & Myers, 2003). Some concerns have also been raised about parent and teacher rater reliability in relation to gender, poverty and race (Bussing et al., 2008). Evidence does, however, suggest that the SNAP-IV (26-item) has acceptable internal consistency and a factor structure consistent with the two-factor solution of ADHD symptoms (Inattention and Hyperactivity; Bussing et al., 2008). While the literature may caution against using the SNAP-IV for diagnostic purposes, the evidence suggests that the SNAP-IV satisfactorily distinguishes children with differing levels of ADHD concern (Bussing et al., 2008) and its use within this thesis is exploratory. As with the RCADS, the SNAP-IV is used within this thesis to confirm the validity of the SDQ by comparing reports of similar constructs i.e. wellbeing and/or behaviour. For example, the adult informed SDQ sub-scale scores will be compared with adult informed SNAP-IV sub-scale scores, adult informed SDQ sub-scale scores will be compared with adult informed SNAP-IV sub-scale scores, and self-reported SDQ sub-scale scores will be compared with adult informed SNAP-IV sub-scale scores.

Chapter Three: Aims of Developing the Tools for this Thesis

In the context of developing the appropriate tools for use in the current thesis, it was important to examine the reliability and validity of self-reported behavioural and emotional difficulties, using the Strengths and Difficulties Questionnaire (SDQ) in 8 to 10 year old schoolchildren. In order to assess the significance of direct information from the child (through their self-report), the scale and direction of differences between parent and child reports (across both sexes, different ages and different constructs) were investigated. This ~~study~~ provided a comprehensive investigation of the complex relationships between informant report and young children's self-report of behaviour and wellbeing in a community sample.

This was achieved by exploring the psychometric properties, reliability and validity of SDQ data collected from different informants; parents, teachers and 8 to 10 year old schoolchildren.

More specifically, the aims of the study were:

1. To explore the ***reliability*** of the SDQ in 8 to 10 year old schoolchildren by studying the internal consistency and inter-rater agreement (between parent, teacher and self-report scores) of the SDQ sub-scales
2. To explore the ***validity*** of the SDQ in 8 to 10 year old schoolchildren by studying the concurrent validity of the SDQ sub-scales (for teacher, parent and self-report). Firstly, by comparing internal construct sub-scale scores on the SDQ (such as Emotional Difficulties and Peer Problems) with the corresponding sub-scale scores on the RCADS, and secondly, by comparing behaviour sub-scale scores on the SDQ (such as Conduct and Hyperactivity) with the corresponding sub-scale scores on the SNAP-IV.
3. To evaluate the proposed 5-scale structure of the SDQ derived from teachers, parents and self-reports using item correlation analysis and factor analysis

3.1 Hypotheses

1. Reports from different informants on the same measure of wellbeing and/or behaviour will be consistent, demonstrating reliability between informants
    a. Self-report, Parent and Teacher scores will correlate significantly on the SDQ

2. Reports of the same construct of wellbeing and/or behaviour will be consistent across different measures, demonstrating validity of the SDQ
    a. Self-reported SDQ sub-scale scores will correlate with corresponding self-reported RCADS sub-scale scores
    b. Self-reported SDQ sub-scale scores will correlate with corresponding parent informed RCADS sub-scale scores
    c. Parent informed SDQ sub-scale scores will correlate with corresponding parent informed SNAP-IV sub-scale scores, and teacher informed SDQ sub-scale scores will correlate with corresponding teacher informed SNAP-IV sub-scale scores
    d. Self-reported SDQ sub-scale scores will correlate with corresponding parent informed SNAP-IV sub-scale scores, and self-reported SDQ sub-scale scores will correlate with corresponding teacher informed SNAP-IV sub-scale scores
3. All versions of the SDQ (self-report, parent and teacher) will replicate the five-factor structure proposed by Goodman (1997) in a sample of young schoolchildren aged 8 to 10 years old.

Chapter Four: The Methodology of Developing the Tools for this Thesis

4.1 Design

This study aimed to present evidence for the reliability and validity of the self-report SDQ in younger children (see Chapter 2). This correlational study employed an observational design to compare the reports of behaviour and emotional difficulties of young children (8 – 10 years-old) from multiple informants (child, parent and teacher) using the SDQ. It also compared the self- report of emotional difficulties using the SDQ with the self-report of difficulties using the RCADS (a self-report measure of anxiety and depression, with established validity in a younger population; see section 2.5.2) and compared the self- report of behavioural difficulties using the SDQ with the parent/teacher- report of difficulties using the SNAP-IV (a reliable and valid measure of behavioural difficulties in children; see section 2.5.3).

The study took place within UK primary schools. The testing of the children was conducted in the school environment, while the testing of the parents and teachers was done via written communication. The questionnaires were distributed and completed in either paper or electronic format.

4.2 Ethical clearance

This study received ethical approval from the School of Psychology Research Ethics Committee (SoPREC) at the University of Leeds. The SoPREC reference number and date of approval are: 17-0253 and 02/10/2017, respectively. All researchers involved in the study obtained enhanced DBS clearance.

4.3 Recruitment

Children aged 8 to 10 years old were recruited to participate in this study. The children were recruited from primary schools in the local area (Leeds) and in Northumberland. The children were recruited from a pool of children from two academic years. Ages 8 to 10 years correspond to compulsory primary school years 4 and 5 in the British School System. Both years correspond to Key Stage 2 in the education system. In total, the study aimed to recruit sufficient children and their parents/teachers, allowing for dropouts (20% attrition), to complete with three sets of ratings (by parents, teachers and the children themselves)

available, in order to perform a factor analysis of the respective responses on the SDQ. Tabachnick and Fidell (2007) suggest having at least 300 cases (pg. 613) but that a sample size of 150 cases is sufficient if solutions have several high loading marker variables (>.80). Appendix 7 shows the flow diagram of recruitment.

4.4 Consent

Before consent was sought, the pupils were told about the research study via an assembly held by the researchers at the schools. Following this assembly, parents/carers received the letter and information sheet (Appendices 8 - 17) sent to their home to inform them about the study. Parent and teacher information sessions were offered at each school for parents/carers/teachers to attend should they have any questions or queries – although this was not taken up. The teaching staff were given detailed explanations of what the study entailed and how their classes would be affected from the head teachers in school meetings.

Consent was gained from parents/carers in the form of opt-in permission letters sent home with the children and assent was obtained from each individual child on each day of testing. Information on any current medication or other reasons which may preclude participation in the study (e.g. severe autism) was obtained from school records via the head teacher, prior to distribution of the opt-in letters.

An honorarium was given for the participation of the child and the parent in the form a ticket for a prize draw (there were two prize draws, one for each region: Leeds and Northumberland). There was the chance to win one £40, two £20, and one £10 in Love2Shop vouchers. To participate the child completed two questionnaires; the SDQ self-report version and the RCADS self-report version, and the parent completed three questionnaires; the SDQ informant version, the RCADS informant version and the SNAP-IV. An honorarium was also given for the participation of the teachers in the form of £15 in Love2Shop vouchers on the completion of two questionnaires (the SDQ informant version and the SNAP-IV) for each child in their class[1].

4.4.1 Ongoing Engagement

Parents/guardians were asked to provide a contact number (when returning their opt-in slip) to allow for phone check-in. Researcher contact details were also provided for information sharing and/or the arrangement of any support in the completion of questionnaires, for

[1] One school chose to use this money to provide teaching cover (to allow time to complete the questionnaires). This was dependent on school preference, as determined by the head-teacher.

example, out of hours (early evening and weekend) home visits. A system was established whereby any persons undertaking a home visit had support mechanisms in place[2].

4.4.2 Study Withdrawal

Participants and their parents/carers were told that they could withdraw participation at any point before or during the study (up until the point of write up – May 2018) without giving a reason, and that all information gathered would be kept strictly confidential and anonymised. No participants withdrew. Researchers withdrew one child's dataset from the analysis due to issues of non-compliance.

4.5 Measures

4.5.1 Screening Measures

Parents/carers were required to complete some screening questions (Appendix 9) based on their child/children regarding demographic information and checks for exclusion criteria, which were incorporated within the parent information sheet/letter.

4.5.2 Outcome Measures

Data on each child was collected by using the SDQ (completed by the teacher, parent and the child themselves), the RCADS (completed by the parent and the child themselves) and the SNAP-IV (completed by the teacher and the parent). These measures are outlined in section 2.5.

4.6 Procedure

4.6.1 Participant Screening

Participants had to meet all of the following inclusion criteria and none of the exclusion criteria to participate in the study. Inclusion criteria: Girls or boys, aged 8 to 10 years old, ability to follow verbal and simple written instructions in English, ability to give written informed consent and ability to understand cognitive written instructions and measures. Exclusion criteria: Learning disabilities which interfere with the ability to understand written or verbal communications, visual impairment that precludes the ability to follow read written instructions and measures, inability to understand the objective of the measures, or complete the items, acute illness, or feeling unwell, within the week prior to testing and

[2] A risk assessment was approved as part of the ethics submission

current administration of any psychotropic, antibiotic medication or supplementation in the month prior to testing, or during testing.

4.6.2 Setting

The self-reported measures were administered during normal school hours on school days (Monday-Friday only). Conducting the study within the school environment aimed to ensure that any support required by the children was on-hand by the researchers. Furthermore, conducting the research in the participant's normal environment and routine alongside their normal lessons offered greater ecological validity and limited potential behaviour change that might occur in a novel environment.

The informant-rated measures were completed outside of school hours, either online (Bristol Online Surveys) or in paper format.

4.6.3 Study Procedure

The study (see Figure 2) ran over a 5 week period; one week for recruitment and the sending home of letters, information sheets and opt-in consent forms. This was followed by a two week gap, to allow 14 days in which participants could decide whether they would like to take part in the study. During this time the screening questionnaires were completed by the parents.

On testing days, children whose parents had opted in were approached to complete the measures. They first read the accessible information sheet and provided assent. They then completed the self-report behaviour and wellbeing measures. A study letter for the parents was then sent home with the children and a study letter for the teachers was distributed to the relevant tutors. These included details on the completion of the informant-rated behaviour and wellbeing measures, which were available online (Bristol Online Surveys) and in paper format. The parents and teachers then had a two-week period to complete the informant-rated behaviour and wellbeing measures.

4.7 Approach to Data Analysis

This section outlines the strategy with respect to case inclusion / exclusion and validity checks of the data. It also presents the analysis strategy including justification for the use of the statistical tests employed. All data were entered into Microsoft Excel and organised for export into SPSS (v22) for statistical analysis.

4.7.1 Data Screening

All missing items were dealt with as per the error protocol (see section 4.7.3)

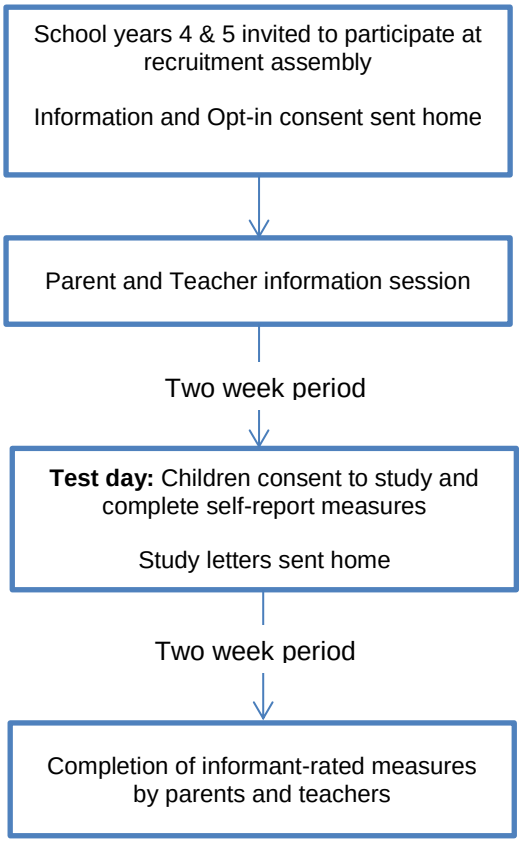


Figure 2: Participant flow chart outlining study design and procedure in each school

To check for outliers, minimum and maximum scores were examined for all dependent variables. This was corroborated by inspection of box plots. Outliers were then checked for feasibility and any data entry errors corrected. Outliers were considered to be values more than three standard deviations from the mean, while data entry errors could be identified as values outside the minimum or maximum score on the Likert scale. Any single data point considered to be an outlier was removed from the analysis.

Normality of variables was assessed graphically, using histograms, and statistically, using values of skewness and kurtosis. The Shapiro-Wilks test was used to check the assumption of normality in the data

For data that fulfilled the normality assumption, ANOVA (Analysis of Variance) analysis was performed whilst the non-parametric equivalent (Kruskal-Wallis test) is presented for data which were not normally distributed. Data were checked for homogeneity of variance using Levine's test. Where this assumption was not met, the adjusted Levene's statistic is presented.

4.7.2 Questionnaire Analysis

4.7.2.1 SDQ

Independent of informant, the 25-item SDQ has five sub-scales (of 5 items each; see section 2.5.1 and Appendix 1,2 and 3) and three composite scores ("Total Difficulty', 'Externalising' and 'Internalising'). The scores on each sub-scale are expressed as the sum of the ratings from the individual items that the sub-scale is composed of. For missing items on the SDQ, scores were scaled up pro-rata if a minimum of 3 items (per sub-scale) had been completed. The composite scores were calculated by summing the relevant sub-scale scores.

4.7.2.2 RCADS

The 47-item (parent informant) and the 25-item (self-report) questionnaires each comprise six sub-scales and two composite scores (see Appendix 1, 4 and 5). They also produce a total anxiety score and a total internalising score. The scores on each sub-scale are expressed as the sum of score from the individual items which make-up the sub-scale. These raw scores are, typically, then given a corresponding t-score for the appropriate age group. A t-score is a common form of standardising test statistics, the t-score formula enables an individual score to be transformed into a standardized one (between 1-100), which helps to compare scores within a population. As the majority of the scores were expected to be below the clinical threshold, raw scores were used for within-informant comparisons. For missing items on the RCADS, scores were calculated pro-rata using the remaining items within a scale. As per recommendations, scales missing more than 2 items were not scored.

4.7.2.3 SNAP-IV

The 18-item, parent and teacher informant, SNAP-IV questionnaire has 2 sub-scales (Appendix 1 and 6). It also produces a total score. Scores are expressed as the average rating per item. For missing items, sub-scale scores were scaled up pro-rata.

4.7.3 Error Protocol for the Management of Missing Data

4.7.3.1 Missing item data

As discussed above, any missing items were dealt with in line with the questionnaire scoring protocol. Where minimum data requirements were met, missing variables were calculated using mean substitution (pro-rata). The missing data points were replaced with the mean of the completed items within that scale. This procedure avoids the removal of the entire variable data set for that participant which would reduce the number of participants and potentially increase variance in the dataset. However, it is important to highlight that this method increases the risk of convergence to the mean and a false positive result.

4.7.3.2. Missing variable data

To avoid removing observations with missing data in a list-wise manner, and thus, the removal of entire participant data sets, pair-wise data deletion was used where possible. Pair-wise deletion, also known as available case analysis, includes all available data for each participant. For example, when conducting a correlation on multiple variables, the statistical package (SPSS in this case) will perform the bivariate correlation between all available data points, and ignore only those missing values if they exist on some variables. In this case and as reported in this thesis, pair-wise deletion results in different sample sizes for each correlation.

4.7.4 Statistical Analysis

4.7.4.1 Descriptive Properties

For the purposes of analysis, the data collected from the different schools was consolidated into one data set. The schools represent a homogeneous sample, as all the schools sampled were semi-rural, suburban schools with primarily white catchment areas. To confirm this, the schools were compared using data collated by the Consumer Data Research Centre (CDRC, Accessed 2018). This included comparisons across classifications relating to the Indices of Multiple Deprivation (IMD), indicators relating to the Access to Healthy Assets & Hazards Index (AHAH) and metrics from the Consensus datasets.

Differences in informant, gender and year group on the SDQ sub-scores were analysed using univariate analysis of variance (ANOVA) to explore effects of the between-subjects factors (informant, gender and school year group) on the dependent variables (SDQ sub-scale scores). Significant main effects for variables were followed up with Bonferroni corrected pairwise comparisons. This post hoc analysis is presented as the mean difference between groups (plus or minus the S.E) followed by the p value.

4.7.4.2 Reliability

To assess the reliability of the SDQ sub-scale scores across informants, Pearson's Product Moment correlation coefficient was calculated for each subscale between participant data provided by teachers, parents and the children themselves. This analysis used pair-wise deletion (as discussed in section 4.7.3.2). Spearman's rho correlation coefficients were also calculated when subscale score variables were skewed. When reporting inter-rater reliability, caution should be exercised in the interpretation of correlations (as the Pearson Product Moment correlation is unaffected by the presence of systematic bias). It is therefore, recommended that both the correlation coefficient and Cohen's kappa be considered to uncover non-random examiner error (Hunt, 1986). Cohen's kappa is a measure of agreement between informants that accounts for agreement based on chance. The

assumption is that data collectors are likely to guess when they are not sure of the correct answer and that a certain number of these guesses are likely to be congruent. Therefore, reliability statistics should account for random agreement. The kappa value can range from -1 to +1.

Cronbach's alpha coefficients were calculated to estimate the internal consistency associated with the items from which a composite score is derived, in order to provide a measure of scale reliability. This analysis used list-wise deletion (see Section 4.7.3.2). Scores yielding a minimum reliability of 0.8 are considered acceptable (Lance, Butts, & Michels, 2006).

4.7.4.3 Validity

To assess the validity of the SDQ sub-scale scores, the concurrent validity between the SDQ sub-scales and the corresponding sub-scales on the RCADS and the SNAP-IV were analysed using Pearson's Product Moment correlation coefficients.

4.7.4.4 Factor Analysis

To investigate the number of constructs and the structure of the SDQ, when completed by the teachers, parents and the children themselves for this sample of young children, an exploratory factor analysis (EFA) was conducted.

To determine the factorability of the dataset, several well-recognised criteria for the factorability of a correlation were used. Initially, it was confirmed that all 25 items of the SDQ correlated at least 0.3 with at least one other item, suggesting reasonable factorability (Tabachnick & Fidell, 2007). Secondly, the overall Kaiser-Meyer-Olkin measure of sampling adequacy for the data and all Kaiser-Meyer-Olkin values for the individual items (found on the diagonals of the anti-image correlation matrix) were confirmed as exceeding the recommended values of 0.57, which is a requirement for a reliable factor analysis (Tabachnick & Fidell, 2007). Thirdly, Bartlett's test of sphericity was checked for significance, indicating that SDQ items do correlate with one another (Tobias & Carlson, 1969). Finally, the communalities of the data which indicate the proportion of variance in each item that can be explained by the remaining factors (also reported as h^2, defined as the sum of squared factor loadings for the variables) were considered. Similarly, extraction communalities indicate the proportion of each variable's variance explained by the retained factors; low values are not considered to be well represented in the common factor space. It was confirmed that all initial and extraction communalities were above 0.3 (Tabachnick & Fidell, 2007).

The number of factors generated in the final solution was determined based upon a number of theoretical criteria; (1) the "eigenvalue greater than 1.00" rule (Kaiser, 1960), (2)

examining the scree plot (Cattell, 1966) and (3) parallel analysis (Hayton, Allen, & Scarpello, 2004; Ledesma & Valero-Mora, 2007). (1) A factor with an eigenvalue of one can account for as much variance as a single variable. The following logic applies: only a factor that can explain the same about of variance as a single variable is worth keeping. (2) A scree plot shows eigenvalues on the y-axis and number of factors on the x-axis, displaying a downward curve. The point at which the curve levels off, known as the elbow, indicates the number of factors that should be generated by the analysis and the elbow represents the point at which further extraction of factors does not substantially increase variance explained. (3) An alternative method to determine the number of factors to extract/retain is parallel analysis. Parallel analysis, uses a Monte Carlo simulation, to generate random correlation matrices based upon the equivalent number of variables and participants. These are then subjected to principal component analysis (PCA) and the average of their eigenvalues is produced. The criterion for factor extraction is where the eigenvalues of the generated factors exceed those of the experimental data.

The teacher, parent and child informant data were then analysed by factor analysis (SPSS procedure *Classify*) using maximum likelihood and a Direct Oblimin rotation. This was forced to give an oblique solution, with a set number of factors informed by the above criteria. The oblique rotation method allows for natural correlations between factors. The analysis, again, used pair-wise deletion (N = Teacher 111, Parent 101, Self-report 126). The oblique rotation was then subsequently justified by factor correlations exceeding 0.32 (Tabachnick & Fidell, 2007), which is indicative of overlapping variance among factors.

Factor solutions with a Heywood case were not considered legitimate (SAS Guide, Accessed 2018). A Heywood case occurs when one or more communality estimate is greater than one. The communalities of the data indicate the proportion of variance in each item that can be explained by the remaining factors. Hence, a large communality indicates that the variable in question will dominate some factor, changing the structure of the factor solution. This can occur when too many factors are extracted or the sample size is too small.

4.7.4.5 Factor Scores

Composite scores were created for the self-report data, in line with the findings from the exploratory factor analysis. A non-refined method of summing scores by factor was chosen. This involves summing raw scores corresponding to all items loading on a factor (Comrey & Lee, 1992). For items yielding a negative factor loading, the raw score of the item is subtracted rather than added to the computation. This method is consistent with the method to calculate the internal consistency reliability of the new factors, using Cronbach's alpha.

This method was chosen, over alternative non-refined and refined methods, as it is one of the simplest ways to estimate factor scores. An evolving pressure of resource utilisation is

the context for the development of measures of behaviour and wellbeing such as intended by this thesis. Therefore, in a cost and time-effective NHS, it is important to hold in mind factors such as; how easily a measure could be implemented into services, the timely feedback needed to inform practice and their use in cross-service comparison (Wolpert et al., 2014; Wolpert, Ford, et al., 2012). This method would also allow for average scores to be computed, retaining the scale metric, if required. Such averaging would allow for easier interpretation and comparisons across factors when there are differing numbers of items per factor. For these reasons a simple weighting method was chosen. A limitation of this method is that all items on a factor are given equal weight, regardless of their loading on the factor, which can limit reliability to some degree. Furthermore, cross-loading items must be considered. In this case, items were considered to load only onto the factor on which the item had the highest factor loading value. The rationale for this is that interpretation may become difficult if items used to compute factor scores are not independent.

Chapter Five: The Results of the Tool Development for this Thesis

This chapter considers the data collected for the purposes of developing appropriate tools for this thesis. It discusses the characteristics of the sample used, a summary of the means and an overview of the response patterns, followed by analysis to investigate the reliability and validity of the questionnaire used within this sample. This includes assessment of the inter-rater reliability; correlations between **self-** report, **parent** and **teacher** report of behavioural and emotional difficulties and inter-rater agreement and consistency for informants using the SDQ. It also includes assessment of the validity of the SDQ by correlating SDQ sub-scale scores with established measures (the RCADS and the SNAP-IV). The final section (section 5.5) assesses the validity of the factor structure of the SDQ when used with the current sample of teachers, parents and young children (aged 8 – 10 years).

5.1 Participant Characteristics

Self-report data were collected from 126 children across six schools, for these children 101 SDQ were completed by parents and 111 SDQ were completed by teachers. The mean age (S.E) of the sample was 111.5 months (0.34), which is equivalent to 9 years and 3.5 months. Table 1 shows the sample size, gender distribution and mean (± S.E) age of children recruited from each school.

Table 1: The number of participants (children) recruited from each school, by gender and year group.

Total N = 126	Total			Year 4		Year 5		Age (months)
	Total	**Boys**	**Girls**	**Boys**	**Girls**	**Boys**	**Girls**	
School A	50	22	28	13	14	9	14	111.42 (1.02)
School B	22	11	11	5	5	6	6	112.(1.33)
School C	10	4	6	4	1	0	5	112.7 (1.34)
School D	5	1	4	1	2	0	2	112.0 (2.59)
School E	25	14	11	8	7	6	4	110.46 (1.37)
School F	10	6	4	4	3	2	1	111.30 (2.01)
Total	126	61	65	67[3]		55		111.5 (0.34)

[3] Information on age and school year group was missing for 4 participants – one female and three males. The following age analysis excludes all data (reports from teachers, parents and the child themselves) for these participants.

A univariate ANOVA revealed that the mean ages of the children were not significantly different across schools (F (5, 122) = 0.23, p > 0.05). For the purposes of further analysis, the data collected from the different schools has been consolidated into one data set (see section 4.7.4.1).

5.2 Descriptive Properties of the SDQ

5.2.1. SDQ Total Difficulty Score as Assessed by Teacher, Parent and Self-report

The means (S.E) of the SDQ Total Difficulty composite scores for children, as rated by their teacher, parent and the child themselves are shown in Figure 3.

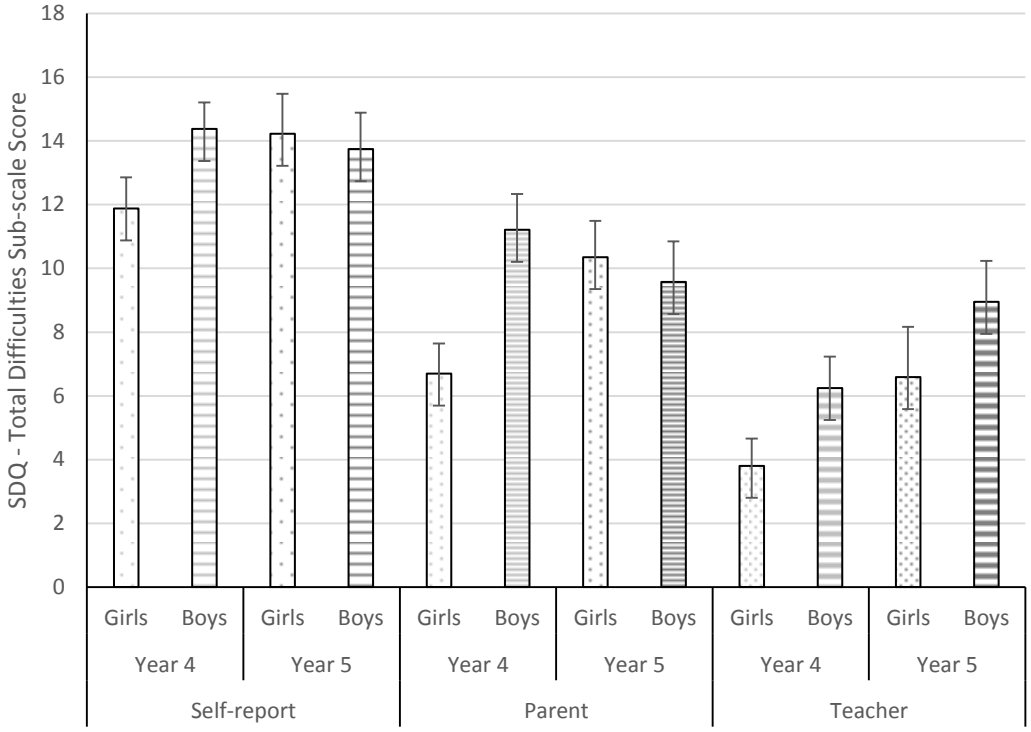


Figure 3: Mean (± S.E) of the SDQ Total Difficulty composite score as rated by each informant (Self-report N = 122; Parent N = 99 and Teacher N = 107), child gender and school year group.

When the results of the SDQ Total Difficulty score were compared, the effects of informant (F (2, 316) = 41.0, p < 0.001), gender (F (1, 316) = 7.03, p < 0.01) and year group (F (1, 316) = 5.36, p < 0.05) were significant.

Follow-up analysis showed that children rated their Total Difficulty higher than their teachers (7.43 (±0.79); p < 0.001) or parents (3.97 (±0.80); p < 0.001). In addition, teachers

rated Total Difficulty lower than parents (-3.46; $p < 0.001$). Boys had higher scores for Total Difficulty than girls (1.76 (±0.66); $p < 0.01$) and the Year 5 children had more difficulties than those in Year 4 (1.54 (±0.66); $p < 0.5$).

The was a significant child gender x year group interaction ($F (1, 316) = 4.39$, $p < 0.05$). In year 4, boys had higher scores for Total Difficulty when compared to girls, but there was no gender difference in year 5 (see Figure 4).

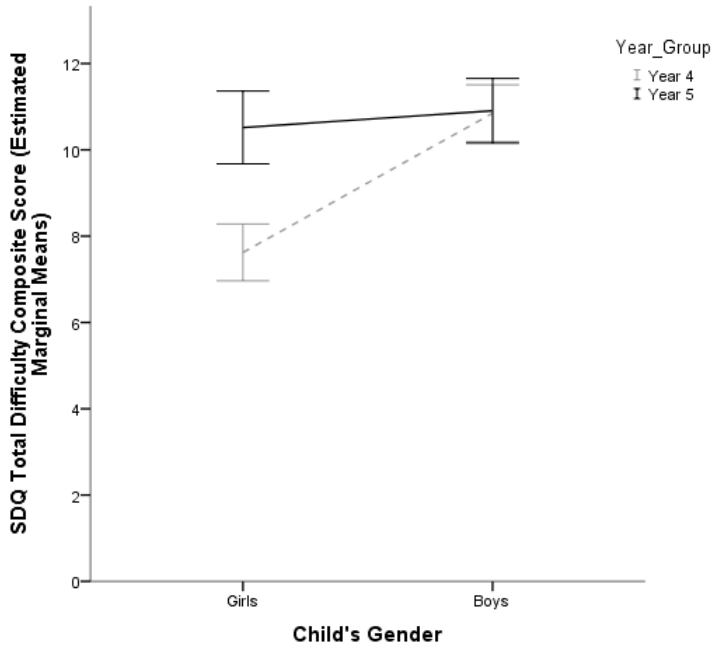


Figure 4: Estimated marginal means of the SDQ Total Difficulty Composite Score, demonstrating the interaction between gender and school year group.

5.2.2 SDQ Sub-scales as assessed by teacher, parent and self-report

The means (S.E) of the primary SDQ sub-scale scores for children, as rated by teachers, parents and the child themselves are shown in Figures 5 (reporting sub-scales for Prosocial Strengths), 6 (reporting sub-scales for Emotional Difficulties and Conduct Difficulties), 7 (reporting sub-scales for Hyperactivity Difficulties and Peer Problems) and 8 (reporting sub-scales for Externalising Difficulties and Internalising Difficulties).

There were significant main effects of informant for all sub-scales; Prosocial ($F(2,316) = 8.79$, $p < 0.001$), Emotion ($F(2,316) = 28.89$, $p < 0.001$), Conduct ($F(2,316) = 32.48$, $p < 0.001$), Hyperactivity ($F(2,316) = 12.85$, $p < 0.001$), Peer problems ($F(2,316) = 26.65$, $p < 0.001$), Externalising ($F(2,316) = 25.13$, $p < 0.001$) and Internalising ($F2,316) = 33.90$, $p < 0.001$). As with the Total Difficulty composite score, self-reported scores were consistently and significantly higher than teacher reported scores (Emotion 2.13 ± 0.30, $p < 0.001$; Conduct 1.86 ± 0.23, $p < 0.001$; Hyperactivity 1.71 ± 0.31, $p < 0.001$; Peer problems 1.72 ±

0.24, $p < 0.001$; External 3.57 ± 0.48, $p < 0.001$ and Internal 3.86 ± 0.47, $p < 0.001$). The exception to this was the Prosocial sub-scale, on which there was no significant difference between scores when informed by the child themselves or the teacher (0.40 ± 0.24, $p = 0.31$). Self-report scores were significantly higher than parent reported scores for all sub-scales (Emotion 1.26 ± 0.31, $p < 0.001$; Conduct 0.97 ± 0.23, $p < 0.001$; Peer problems 1.44 ± 0.25, $p < 0.001$, External 1.27 ± 0.49, $p < 0.05$ and Internal 2.70 ± 0.47, $p < 0.05$), except Hyperactivity (no difference; 0.30 ± 0.32, $p = 1.00$) and Prosocial (scores were higher when informed by the parent than by the child themselves; -0.60 ± 0.25, $p < 0.05$).

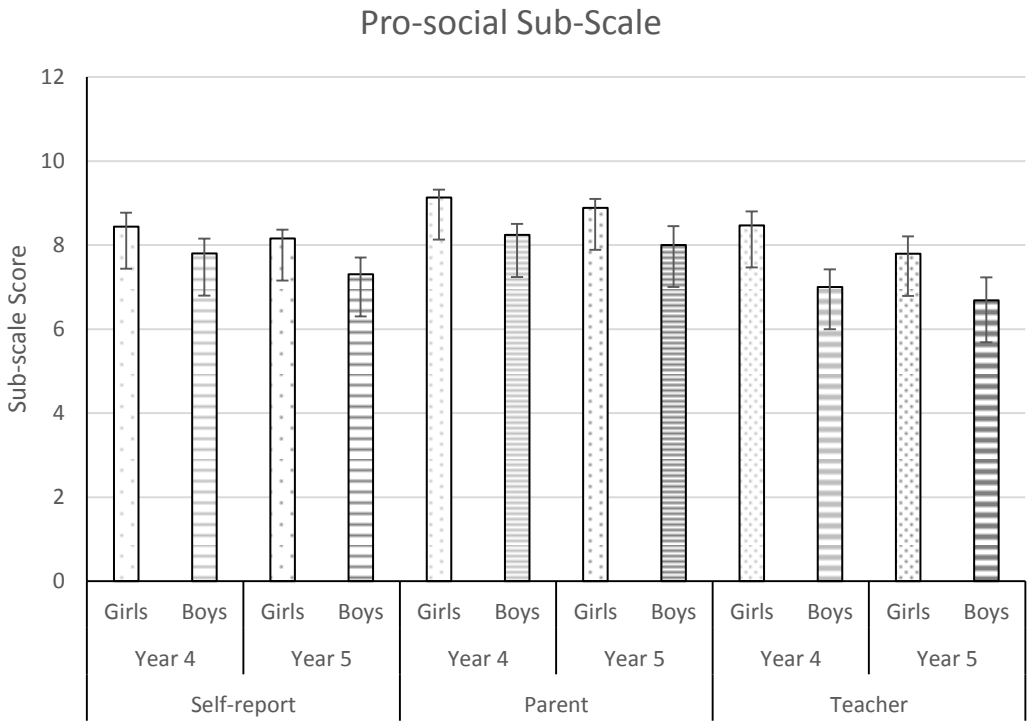


Figure 5: Mean (± S.E) of the SDQ Prosocial sub-scale scores

as rated by each informant (Self-report N = 122; Parent N = 99 and Teacher N = 111), by child gender and school year group.

Parent reported scores were significantly higher than teacher reported scores for all sub-scales (Emotion 0.87 ± 0.31, $p < 0.05$, Conduct 0.89 ± 0.32, $p \leq 0.001$, Hyperactivity 1.41 ± 0.33, $p < 0.001$ and External 2.30 ± 0.50, $p < 0.001$), except for Peer Problems (no difference; 0.28 ± 0.25, $p = 0.80$), Internalised Difficulties (no difference; 1.15 ± 0.49, $p = 0.06$) and Prosocial Strengths sub-scale (scores were non-significantly lower when informed by the parent than by the teacher; -0.40 ± 0.24, $p = 0.31$), although this difference was not significant.

There was a significant main effect of child gender for scores on the Hyperactivity sub-scale ($F(1,316) = 24.04$, $p < 0.001$) and Externalised Difficulties sub-scale ($F(1,316) = 15.78$, $p <$

0.001. Girls scored lower than boys on the Hyperactivity sub-scale (-1.30 ± 0.66, $p < 0.001$) and the Externalised Difficulties sub-scale (-1.60 ± 0.40, $p < 0.001$) when data for these sub-scales were pooled across all informants and school year groups.

There was a significant main effect of year group for scores on the Conduct sub-scale ($F(1, 316) = 8.30$, $p < 0.01$) and the Externalised Difficulties sub-scale ($F(1,316) = 10.00$, $p < 0.01$). Scores for older children (year 5) were higher than corresponding scores for younger children (year 4) on the Conduct sub-scale (0.56 ± 0.19, $p < 0.01$) and the Externalised Difficulties sub-scale (1.27 ± 0.40, $p < 0.01$), when data for these sub-scales were pooled across all informants and child gender. There were no significant interaction effects.

5.2.3 Item level Responses and Descriptive Statistics

Analysis of the responses (see Table 2) at the item level showed relatively low (between 2.5% and 19.2%) endorsement of "certainly true" for all items, except for those within the Prosocial Strengths sub-scale (where endorsements for "certainly true" ranged from 57.1% to 73.6%). The self-report responses had a higher endorsement for the middle response ("somewhat true"; Self-Report 36.6% vs. Parent 29.2% and Teacher 22.8%). Hyperactivity items were rated as "somewhat-" or "certainly true" more frequently than the other difficulties, across all type-of-informants.

The mean (± S.E) scores for each item ranged from 0.13±0.04 to 1.81±0.04 with scores on Item 22 (Steals) and Item 17. (Kind to kids) yielding the lowest and the highest mean scores respectively. Across all SDQ items, skewness ranged from -2.57 to 3.46 and kurtosis ranged from -1.41 to 11.58.

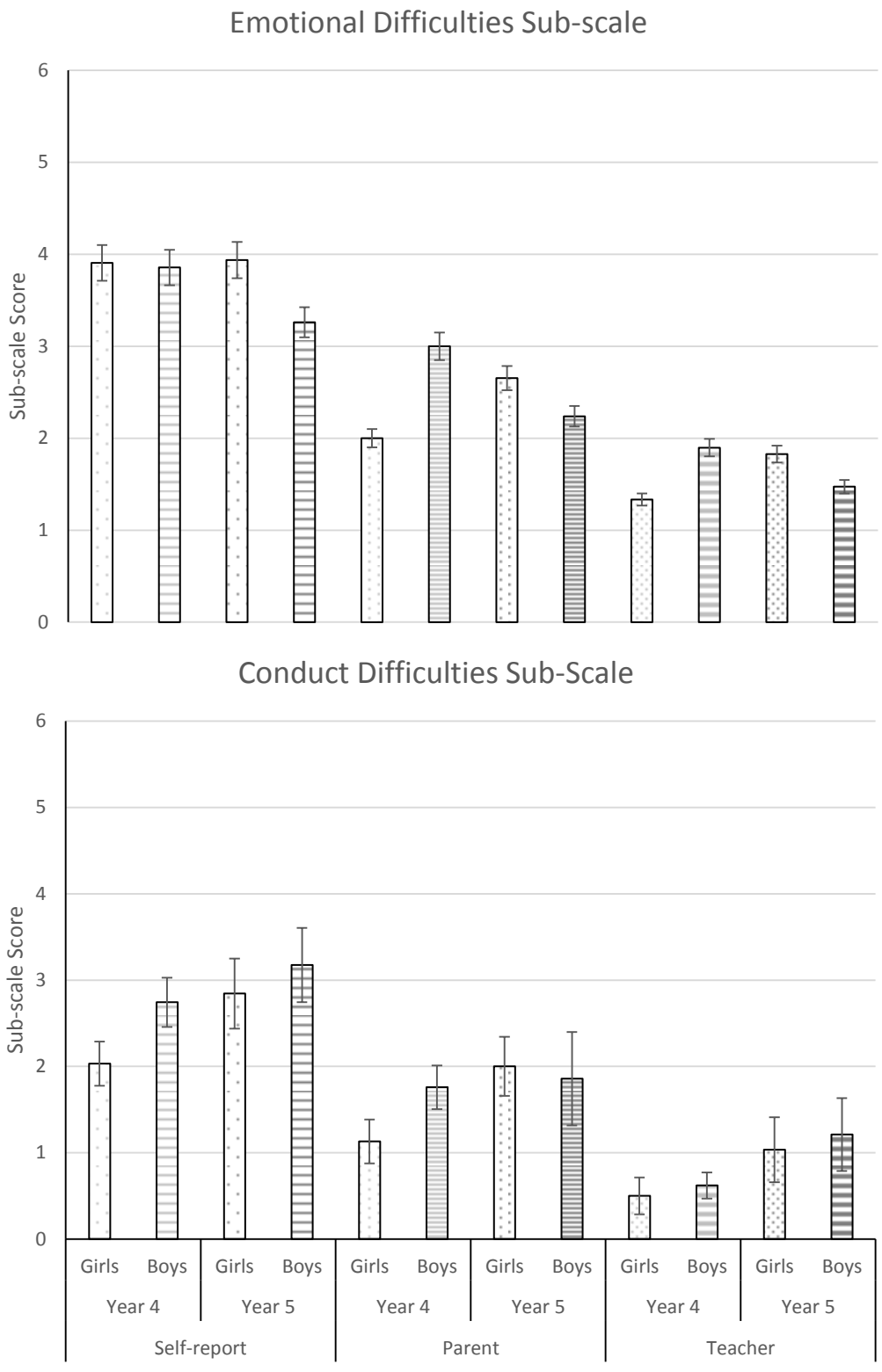


Figure 6: Mean (± S.E) of the SDQ sub-scale scores (1) Emotional Difficulties and (2) Conduct Difficulties

as rated by each informant (Self-report N = 122; Parent N = 99 and Teacher N = 111), by child gender and school year group.

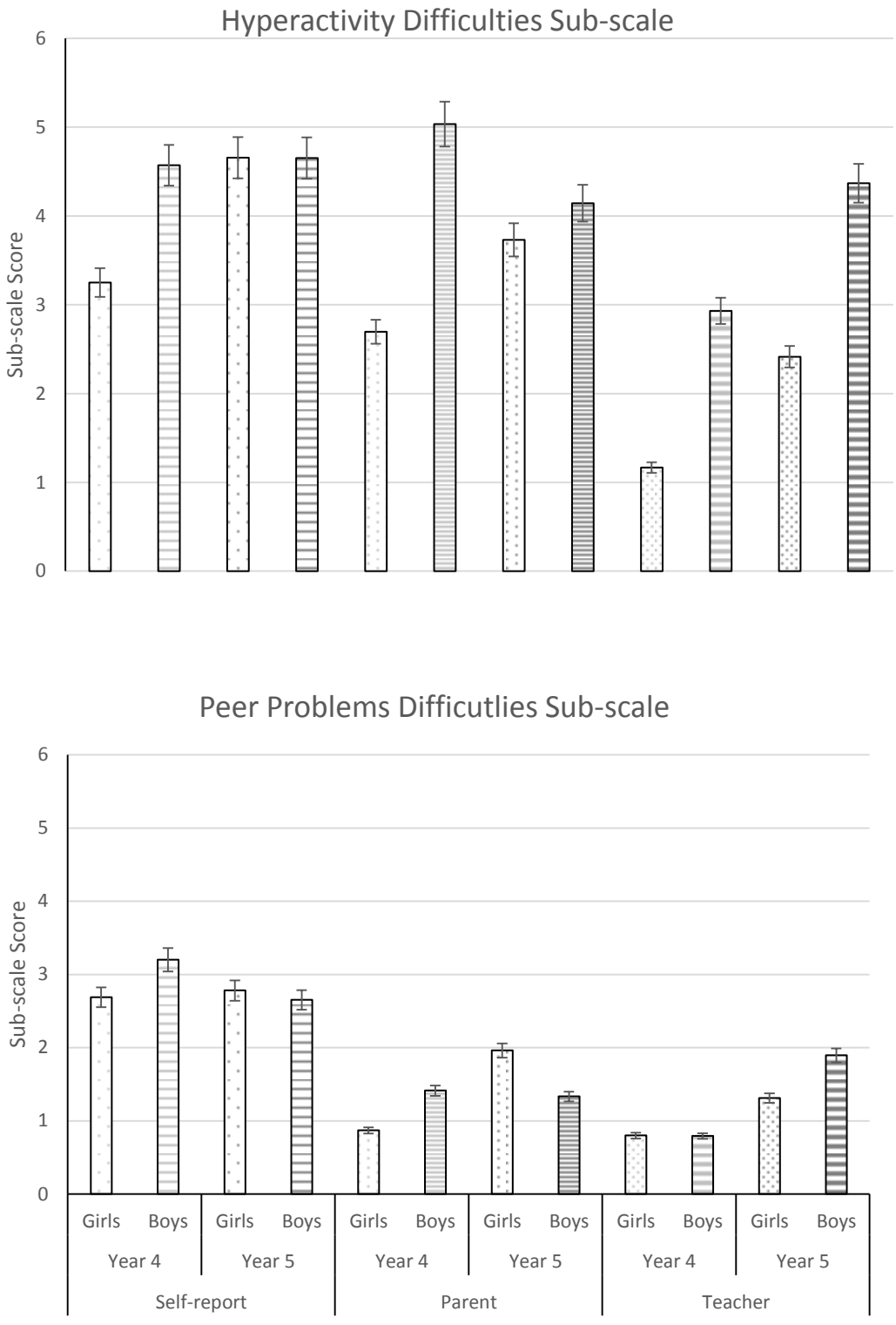


Figure 7: Mean (± S.E) of the SDQ sub-scale scores (1) Hyperactivity Difficulties and (2) Peer Problems

as rated by each informant (Self-report N = 122; Parent N = 99 and Teacher N = 111), by child gender and school year group.

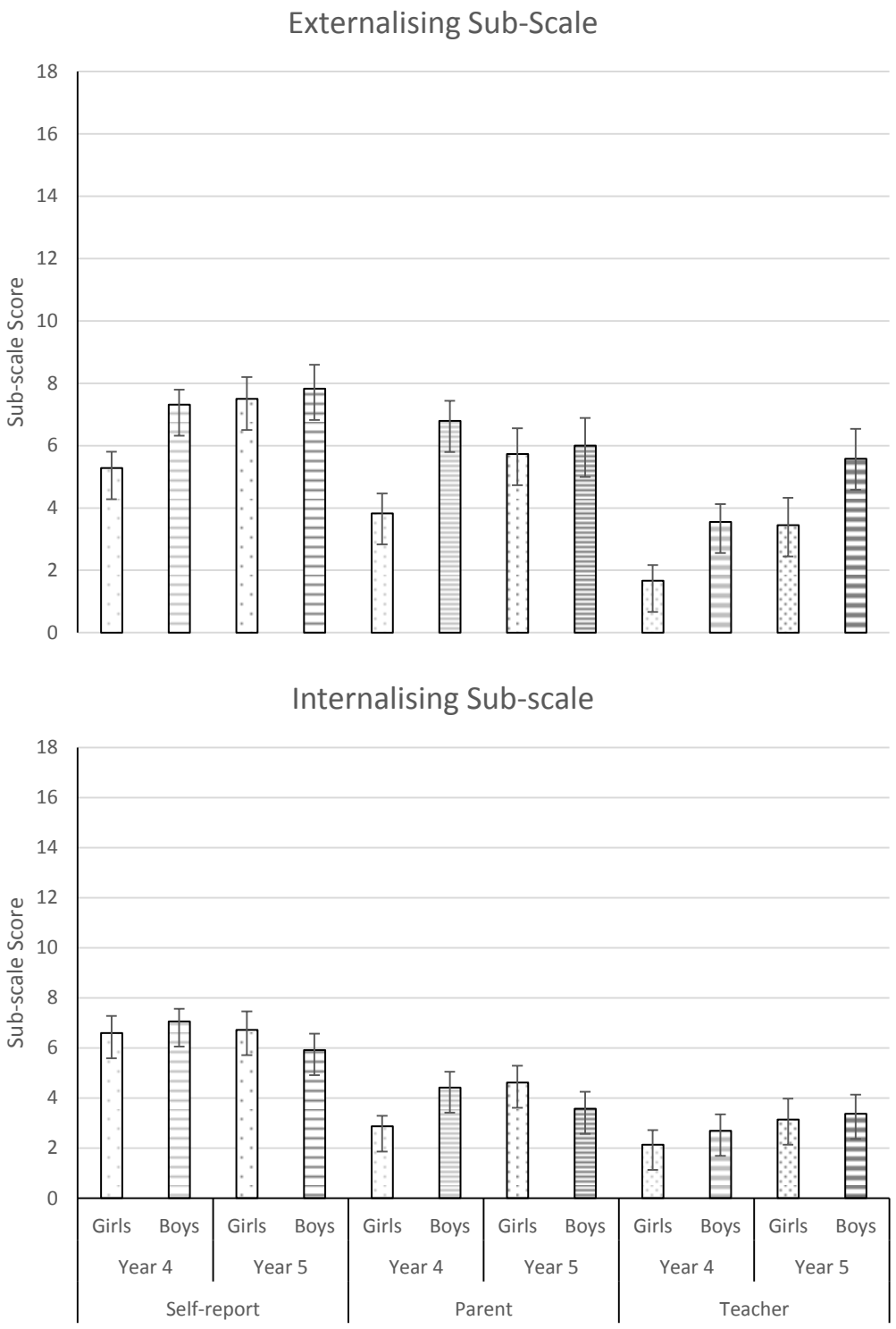


Figure 8: Mean (± S.E) of the SDQ sub-scale scores (1) Externalising Difficulties and (2) Internalising Difficulties

as rated by each informant (Self-report N = 122; Parent N = 99 and Teacher N = 111), by child gender and school year group.

.

Scale	Item #	Item Coding	Item wording	Response Rates (%)								
				Self-Report			Parent-Informant			Teacher-Informant		
				Not True	Somewhat True	Certainly True	Not True	Somewhat True	Certainly True	Not True	Somewhat True	Certainly True
Prosocial	1	Considerate	I try to be nice to other people. I care about their feelings.	0.8	22.2	77	1.0	32.7	66.3	1.8	35.1	63.1
	4	Shares	I usually share with others (food, games, pens)	13.6	38.4	48	2.0	33.7	64.4	6.3	38.7	55.0
	9	Caring	I am helpful if someone is hurt, upset or feeling ill.	5.6	26.2	68.3	2.0	12.9	85.1	7.2	39.6	53.2
	17	Kind to kids	I am kind to younger children.	4	11.2	84.8	1.0	11.9	87.1	1.8	30.9	67.3
	20	Helps out	I often volunteer to help others (parents, teachers, children).	8.7	31.7	59.5	4.0	30.7	65.3	9.0	44.1	46.8
Emotion	3	Somatic	I get a lot of headaches, stomach-aches or sickness.	45.5	42.3	12.2	64.0	27.0	9.0	73.0	20.7	6.3
	8	Worries	I worry a lot.	38.4	42.4	19.2	52.5	37.6	9.9	57.7	30.6	11.7
	13	Unhappy	I am often unhappy, down-hearted or tearful	49.2	43.7	7.1	74.0	21.0	5.0	79.3	14.4	6.3
	16	Clingy	I am nervous in new situations. I easily lose confidence	31.7	37.3	31	44.6	44.6	10.9	77.5	18.9	3.6
	24	Fears	I have many fears. I am easily scared	41.3	36.5	22.2	60.7	29.7	9.9	80.2	13.5	6.3
Conduct	5	Tempers	I get very angry and often lose my temper	31	43.7	25.4	50.0	31.0	19.0	87.4	9.0	3.6
	7a	Obedient	I usually do as I am told.	42.1	50.8	7.1	51.5	45.5	3.0	69.4	27.0	3.6

Table 2: Response rates of the SDQ.

a. Indicates reverse or positively worded items. NB: Shading indicates response preference

	12	Fights	I fight a lot. I can make other people do what I want	71.4	24.6	4	91.1	6.9	2.0	87.3	9.1	3.6
	18	Lies	I am often accused of lying or cheating	52	37.6	10.4	70.3	25.7	4.0	91.0	7.2	1.8
	22	Steals	I take things that are not mine from home, school or elsewhere	90.5	6.3	3.2	96.0	3.0	1.0	98.2	1.8	0.0
Hyper	2	Restless	I am restless, I cannot stay still for long	22.2	47.6	30.2	41.6	42.6	15.8	65.8	22.5	11.7
	10	Fidgety	I am constantly fidgeting or squirming	33.3	40.5	26.2	48.5	34.7	16.8	70.3	20.7	9.0
	15	Distractible	I am easily distracted. I find it difficult to concentrate	27.8	48.4	23.8	35.0	40.0	25.0	47.7	34.2	18.0
	21a	Reflective	I think before I do things	35.7	56.3	7.9	21.8	63.4	14.9	47.3	39.1	13.6
	25a	Persistent	I finish the work I'm doing. My attention is good	40.5	51.6	7.9	31.7	52.5	15.8	52.3	30.6	17.1
Peer	6	Solitary	I am usually on my own. I generally play alone or keep to myself.	51.6	35.7	12.7	68.3	24.8	6.9	85.5	10.9	3.6
	11a	Good friend	I have one good friend or more	84.9	10.3	4.8	89.1	8.9	2.0	83.8	10.8	5.4
	14a	Popular	Other people my age generally like me	41.6	49.6	8.8	79.2	18.8	2.0	70.6	26.6	2.8
	19	Bullied	Other children or young people pick on me or bully me	56.8	29.6	13.6	74.3	23.8	2.0	91.0	8.1	0.9
	23	Better with adults	I get on better with adults than with people my own age	34.1	50.8	15.1	68.3	25.7	5.9	71.2	25.2	3.6

Table 2 continued: Response rates of the SDQ.

a. Indicates reversed or positively worded items, which have been negatively scored. NB: Shading indicates response preference

5.3 Reliability of the SDQ

5.3.1 Inter-rater Correlations

There were a number of significant correlations between the parent and teacher responses (Table 3). However, there are a large number of non-significant correlations between other informant pairs. The children's self-report produced higher total difficulty scores (reflecting higher scores on most of the difficulty sub-scales) as compared to both the parent and teacher informant. The lowest scores were consistently from the teacher informants. The parent informants produced the highest prosocial score. This is reflected in the lack of significant correlations between the children's self-report and the adult informants. The exception to this is the significant correlation between the child themselves and the teacher for the prosocial sub-scale.

Table 3: Inter-rater correlations between informants for each SDQ Sub-scale
Pearson's product moment correlation coefficients * <0.05, ** <0.001

Sub-scale	Self - Parent N 100[4]	Self - Teacher N 111	Parent - Teacher N 89
Prosocial	0.05	-0.18	0.20
Emotion	-0.02	0.01	0.36**
Conduct	0.09	-0.01	-0.06
Hyperactivity	0.10	-0.01	0.43**
Peer Problems	0.01	-0.08	0.37**
Composite Score			
Total Difficulty	0.15	0.01	0.34**
External Difficulties	0.14	-0.05	0.27**
Internal Difficulties	-0.01	-0.00	0.40**

Overall, Table 3 shows that there was most agreement between the ratings of parents and teachers.

5.3.2 Inter-rater Agreement

Table 4 presents the inter-rater agreement for the SDQ scores on participants provided by self-report, parents and teachers. After accounting for chance, there was minimal inter-rater

[4] This value represents the number of available data points for the correlation, missing variable data from informants was excluded in a list-wise manner

agreement across the sub-scales. Inter-rater agreement was only significant for self-report and parent responses on the Emotional Difficulties sub-scale.

Table 4: Cohen's Kappa inter-rater agreement between informant for each SDQ sub-scale

* <0.05

Sub-scale	Self - Parent N 100	Self - Teacher N 111	Parent - Teacher N 89
Prosocial	0.00	-0.04	0.01
Emotion	0.08*	-0.01	0.08
Conduct	0.02	-0.05	0.06
Hyperactivity	0.01	0.02	0.05
Peer Problems	-0.01	0.06	0.05
Composite Score			
Total Difficulty	0.01	0.01	-0.03
External Difficulties	-0.03	0.00	0.00
Internal Difficulties	0.01	-0.02	0.04

5.3.3 Internal Consistency

As informants, the teachers had the best internal consistency, followed by parents, with the lowest internal consistency provided by self-report data. Table 5 (below) presents the internal consistency for the SDQ composite scales calculated for participants as rated by self-report, parents and teachers.

5.4 Validity of the SDQ

5.4.1 Convergent / Concurrent Validity

It was predicted that the SDQ sub-scale scores would correlate with measures of similar constructs using other tools. The concurrent validity was assessed by comparing the total difficulty scores and the sub-scores for the corresponding domains of the SDQ and the RCADS using Pearson's product moment correlation coefficient (Table 6).

Table 5: Cronbach's alpha coefficients for informants on each SDQ sub-scale.

Where missing data occurred, this analysis used list-wise deletion (see section 4.7.4.2) * < 0.80

Sub-scale	Number of items in sub-scale	Self-Report	Parent	Teacher
Prosocial	5	0.57	0.57	0.82*
Emotion	5	0.65	0.62	0.88*
Conduct	5	0.60	0.65	0.76
Hyperactivity	5	0.61	0.81*	0.89*
Peer problems	5	0.49	0.72	0.80*
Composite Score				
Total difficulty	20	0.78	0.82*	0.90**
Externalising total	10	0.71	0.81*	0.88*
Internalising total	10	0.70	0.72	0.88*
Mean of internal consistency of sub-scales		0.64	0.64	0.76

Teacher informants did not complete the RCADS. For the parent/parent comparisons, the correlations between sub-scale scores on the SDQ and corresponding sub-scale scores on the RCADS were in the range of 0.10 – 0.59. The highest correlations were between the SDQ Emotional Difficulties sub-scale and the RCADS Total Difficulty sub-scale and between the SDQ Internalised Difficulties composite score and the RCADS Depression sub-scale.

For the self-report/parent comparisons, the correlations between sub-scale scores on the SDQ and corresponding sub-scale scores on the RCADS were in the range 0.02 to 0.21. The highest correlations showed a negative association between the SDQ Prosocial Strengths sub-scale and the RCADS Total Difficulty sub-scale and the RCADS Depression sub-scale.

The self-report/self-report comparisons showed that the self-report SDQ corresponding sub-scales did not significantly correlate with corresponding sub-scales on the self-report RCADS.

Table 6: The concurrent validity of the SDQ, using the RCADS

Pearson's product moment correlation coefficients between parent rated SDQ sub-scale scores and corresponding parent rated sub-scale scores on the RCADS and correlations between self-reported SDQ sub-scale scores and corresponding self-reported sub-scale scores on the RCADS.

NB: An anxiety score is not calculated for the parent version of RCADS. Emboldened comparisons indicate the significant correlations. * = < 0.05. ** < 0.01

Scores SDQ / RCADS	Self-Report / Self-report	Parent / Parent	Self-Report / Parent
Prosocial / Total	-0.01	-0.10	**-0.21***
Emotion / Total	0.07	**0.59****	-0.06
Conduct / Total	0.05	0.23*	0.15
Hyperactivity / Total	0.07	0.36**	0.07
Peer Problems / Total	0.02	0.20*	0.10
Total Difficulty / Total	**0.08**	0.43**	0.08
External / Total	0.07	0.26**	0.12
Internal / Total	0.06	0.50**	0.02
Prosocial / Depression	-0.01	-0.35**	**-0.24***
Emotion / Depression	0.04	0.49**	-0.15
Conduct / Depression	0.05	0.42**	0.19
Hyperactivity / Depression	0.10	0.33**	0.04
Peer Problems / Depression	-0.02	0.34**	0.11
Total Difficulty / Depression	0.06	0.55**	0.05
External / Depression	**0.09**	0.42**	0.12
Internal / Depression	0.02	**0.52****	-0.03
Prosocial / Anxiety	-0.00		
Emotion / Anxiety	**0.08**		
Conduct / Anxiety	0.05		
Hyperactivity / Anxiety	0.04		
Peer Problems / Anxiety	0.03		
Total Difficulty / Anxiety	0.07		
External / Anxiety	0.05		
Internal / Anxiety	**0.07**		

Table 7: The concurrent validity of the SDQ, using the SNAP-IV

Pearson's product moment correlation coefficients between self-report SDQ sub-scale scores and corresponding parent and teacher rated sub-scale scores on the SNAP-IV. In addition to, Pearson's product moment correlation coefficients between parent rated SDQ sub-scale scores and corresponding parent rated sub-scale scores on the SNAP-IV and correlations between teacher rated SDQ sub-scale scores and corresponding teacher rated sub-scale scores on the SNAP-IV.
NB: Emboldened comparisons indicate the significant correlations. * = < 0.05. ** < 0.01

	Inter-rater		Intra-rater	
Scores SDQ / SNAP-IV	**Self-report / Parent**	**Self-report / Teacher**	**Parent / Parent**	**Teacher / Teacher**
Prosocial / Inattention	-0.09	**0.15**	-0.27**	-0.58**
Emotion / Inattention	-0.06	0.11	0.38**	0.38**
Conduct / Inattention	0.07	-0.03	0.32**	0.55**
Hyperactivity / Inattention	**0.12**	0.01	**0.61****	**0.89****
Peer Problems / Inattention	0.05	0.06	0.13	0.39**
Total Difficulty / Inattention	0.06	0.06	0.55**	0.77**
External / Inattention	0.11	-0.01	0.56**	0.83**
Internal / Inattention	-0.02	0.10	0.33**	0.44**
Prosocial / Hyperactivity	-0.05	0.10	-0.26**	-0.38**
Emotion / Hyperactivity	-0.03	**-0.11**	0.41**	0.19
Conduct / Hyperactivity	**0.13**	-0.10	0.56**	0.61**
Hyperactivity / Hyperactivity	0.03	-0.05	0.63**	0.68**
Peer Problems / Hyperactivity	0.05	-0.01	0.16	0.35**
Total Difficulty / Hyperactivity	0.06	-0.10	0.65**	0.61**
External / Hyperactivity	0.08	-0.09	**0.68****	**0.71****
Internal / Hyperactivity	0.01	-0.07	0.36**	0.29**

Concurrent validity was also assessed by comparing the self-reported SDQ Total Difficulty scores and SDQ sub-scale scores with the parent and teacher informed sub-scale scores of the SNAP-IV, using Pearson's product moment correlation coefficients. Further comparison of the parent and teacher reported SDQ total difficulty scores and SDQ sub-scale scores with the parent and teacher reported sub-scales on the SNAP-IV was undertaken, to demonstrate intra-rater concurrent validity (Table 7)

The self-reported SDQ sub-scale scores did not correlate with parent rated Inattention or Hyperactivity sub-scales from the SNAP-IV. Interestingly, the self-reported Hyperactivity sub-scale on the SDQ was most similar to the parent ratings of Inattention on the SNAP-IV (compared to parent or teacher reported hyperactivity), while the self-reported Conduct sub-scale on the SDQ was most similar to the parent ratings of hyperactivity on the SNAP-IV (compared to parent or teacher reported inattention). Of the self-reported SDQ sub-scales, only the prosocial sub-scale on the SDQ correlated with the Inattention teacher-rated sub-scale from the SNAP-IV. None of the self-reported SDQ sub-scales correlated with the teacher rated Hyperactivity sub-scale from the SNAP-IV.

Both parent and teacher reports on the SDQ sub-scales correlated with their Inattention and Hyperactivity sub-scores from the SNAP-IV. The SDQ Hyperactivity sub-scale and the SNAP-IV Inattention scale correlated strongly, across both the teacher and parent ratings. The SDQ External Difficulties sub-scale correlated highly with the SNAP Hyperactivity scale across both the teacher and parent ratings.

5.5 Factor Analysis of the SDQ Data

Three factor analyses were run according to informant (teacher, parent and the children themselves). The teacher data is presented first to confirm previous findings, whilst the self-report data is presented last.

5.5.1 Factor Analysis of SDQ Teacher informant data

5.5.1.1 Assumption Testing for Teacher Informant SDQ Data

Initially, the factorability of the dataset was examined (see section 4.7.4.4). It was observed that all 25 items of the SDQ correlated at least 0.3 with at least one other item. Secondly, the overall Kaiser-Meyer-Olkin measure of sampling adequacy for the data was 0.86 and all Kaiser-Meyer-Olkin values for the individual items were greater than 0.72. Bartlett's test of sphericity was highly significant ($\chi^2(300)=1607.56$; $p<0.001$). Finally, only one item showed an extraction communality value of less than 0.2 (Item 22. "Steals"; 0.16).

These preliminary analyses confirm the appropriateness of factor analysis for this data set. No items were removed.

5.5.1.2 Number of Factors Extracted from Teacher Informant SDQ data

For the teacher-informant data, the scree plot shows two inflection points: one at factor 3 and the other, less apparent, at factor 6. The factors to the left of the inflection point (also known as the 'elbow'), i.e. those with the 5 largest eigenvalues, were retained. Figure 9 shows that, using the "eigenvalue greater than 1.00" rule, a 5 factor solution would also be generated.

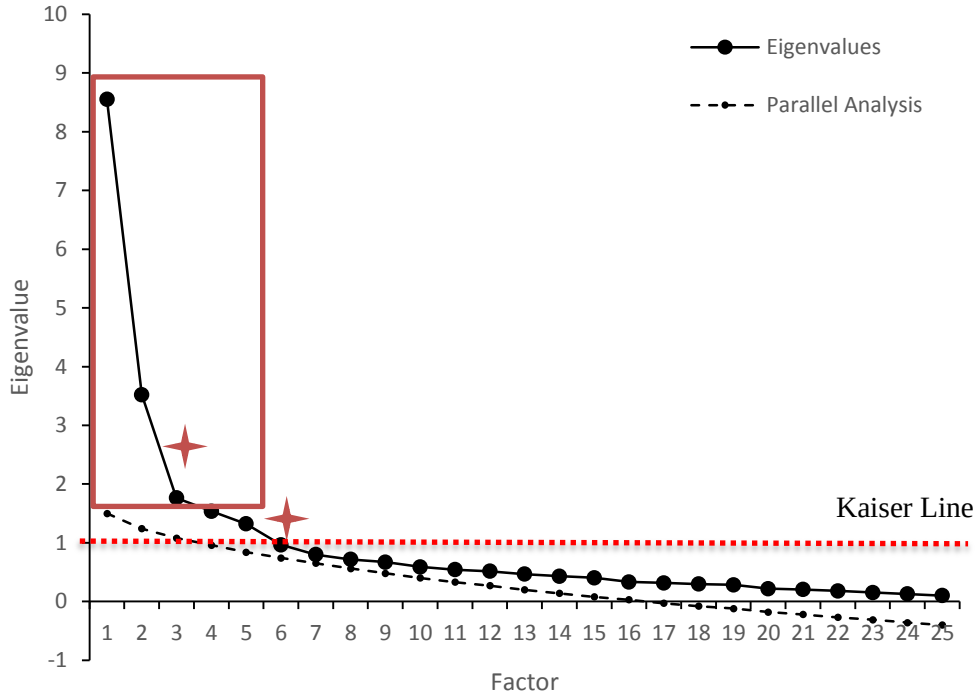


Figure 9: Scree plot and parallel analysis of eigenvalues for the teacher informant data. ✦ indicates the inflection points.

The Monte Carlo Simulation (see section 4.7.4.4) was run for the 105 eligible cases, with 25 variables, at a significance of 0.95, computing 1000 data sets, for common factor analysis using permutations of the raw data set. The matrix identified five significant factors.

Confirmed by the parallel analysis, the "eigenvalue greater than 1.00" rule and the scree plot, a five-factor solution was generated. The Goodness-of-fit test for the 5 factor model ($\chi^2(185) = 229$, $p < 0.05$) indicated some remaining variance to be accounted for within the residual correlation matrix.

Table 8: Factor loadings and communalities for the teacher informed response data

Based on a maximum likelihood factor analysis with a direct oblimin rotation for 25 items from the Strengths and Difficulties Questionnaire (SDQ; N = 111). Note: Correlations below 0.3 are not shown. Red text indicate items that load on multiple factors. Key: H = Hyperactivity, E = Emotional Difficulties, PP = Peer Problems, C = Conduct Difficulties and PS = Prosocial Strengths.

Item / SDQ Sub-scale	New Factors					Communality
	Hyperactivity	Emotional Difficulties	Peer Problems	Propriety[5]	Prosocial Strengths	
15. DISTRACTABLE_H	0.81					**0.81**
10. FIDGETY_H	0.76					**0.64**
25. PERSISTENT(R)_H	0.71					**0.80**
2. RESTLESS	0.67					**0.67**
21. REFLECTIVE(R)_H	0.45				-0.40	**0.71**
16. CLINGY_E		0.87				**0.79**
24. FEARS_E		0.86				**0.70**
8. WORRIES_E		0.75				**0.72**
13. UNHAPPY_E		0.70				**0.72**
3. SOMATIC_E		0.39				**0.33**
6. SOLITARY_PP			0.75			**0.55**
11. GOOD FRIEND(R)_PP			0.72			**0.64**
23. BETTER WITH ADULTS_PP			0.54			**0.42**
19. BULLIED_PP			0.54			**0.49**
14. POPULAR(R)_PP			0.43	-0.33	-0.35	**0.70**
12. FIGHTS_C				-0.79		**0.61**
5. TEMPERS_C				-0.67		**0.58**
18. LIES_C				-0.64		**0.62**
7. OBEDIENT(R)_C	0.39			-0.44		**0.57**
4. SHARES_PS				0.38	0.33	**0.55**
22. STEALS_C				-0.36		**0.31**
20. HELPFUL_PS					0.72	**0.52**
17. KIND_PS				0.32	0.63	**0.55**
9. CARING_PS					0.53	**0.52**
1. CONSIDERATE_PS				0.40	0.43	**0.72**
Eigenvalue	**8.55**	**3.52**	**1.76**	**1.54**	**1.32**	
% of total variance	**34.21**	**14.09**	**7.05**	**6.15**	**5.29**	
Total variance					**66.79%**	

[5] See section 6.2.4.1 for details on factor labelling

5.5.1.3 Factor Loadings for Teacher Informant SDQ Data

The factor structure proposed by Goodman (1994) appeared to map onto extracted factors and was retained: (1) Hyperactivity Difficulties, (2) Emotion Difficulties, (3) Peer Problems, (4) Conduct Difficulties and (5) Prosocial Strengths. Table 8 presents the final pattern matrix (after rotation) and factor loadings of each item of the SDQ. The majority of the factor loadings of the individual items were highly significant. The lowest was 0.36 (Item 22. "Steals").

None of the items were eliminated because they did not contribute to a simple factor structure or failed to meet a minimum criteria of (a) having a primary factor loading of 0.40 or above, and (b) no cross-loading of 0.30 or above. For example, Item 21 (Reflective) had a positive factor loading of 0.45 on Hyperactivity and a negative factor loading of -0.40 on Prosocial. Item 14 (Popular) had a positive primary factor loading of 0.43 on Peer Problems and negative cross-loadings of 0.33 and 0.35 on Conduct and Prosocial, respectively. Item 7 (Obedient) had a negative primary factor loadings of 0.44 on Conduct and a positive cross-loading of 0.39 on Hyperactivity. Items 14 (Popular), 4 (Shares), 17 (Kind) and 1 (Considerate) all had similar positive factor loadings between 0.32 – 0.40 on both Conduct and Prosocial.

5.5.1.4 Factor Correlations for Teacher Informant SDQ Data

The factor correlation matrix (see Table 9) showed moderate positive correlations between factor 1 (Hyperactivity) and both factor: 4 (Conduct) and with factor 5 (Prosocial). There was also moderate positive correlation between factor 2 (Emotion) and factor 3 (Peer Problems).

Table 9: Factor correlation matrix for teacher informant data, based on a five-factor oblique rotation.

Factor	Hyperactivity	Emotional Problems	Peer Problems	**Propriety**[6]	Prosocial Strengths
1		0.22	0.26	-0.46*	-0.40*
2			0.38*	-0.21	-0.01
3				-0.22	-0.10
4					0.29
5					

[6] See section 6.2.4.2 for details on factor labelling

5.5.2 Factor Analysis of SDQ Parent informant

5.5.2.1 Assumption Testing for Parent Informant SDQ Data

The factorability of the dataset was examined (see section 4.7.4.4). All but one (Item 3. "Somatic") of the 25 items correlated at least 0.3 with at least one other item. The overall Kaiser-Meyer-Olkin measure of sampling adequacy was 0.71 and all Kaiser-Meyer-Olkin values for the individual items were greater than 0.57. Bartlett's test of sphericity was highly significant ($\chi^2(300)=868.19$; $p<0.001$). Three variables showed extraction communality values of less than 0.2 (Item 3. "Somatic" = 0.07; Item 4. "Shares" = 0.16; and Item 20. "Helpful" = 0.16) but were retained in the factor analysis to allow comparisons with the factor analysis of the teacher informant and self report.

Overall, these preliminary analyses confirm the appropriateness of factor analysis with this data set.

5.5.2.2 Number of Factors for Parent Informant SDQ Data

Figure 10 shows that, the "eigenvalue greater than 1.00" rule, indicated an 8 factor solution should be generated. The initial eigenvalues indicated that the first 5 factors explained 21%, 10%, 9%, 7% and 5% of the variance respectively. The sixth, seventh and eighth factors each explained between 4.2 – 4.9% of the variance.

The scree plot (Figure 10) appears to have two inflection points: one at factor 2 and the other at factor 5. The factors to the left of the inflection point were retained, i.e. the 4 largest eigenvalues.

Parallel analysis using a Monte Carlo simulation was employed to assess the number of factors retained. The analysis was run for the 97 cases, with 25 variables, at a significance of 0.95, computing 1000 data sets, for common factor analysis using permutations of the raw data set. The matrix identified **four** significant factors (Table 11).

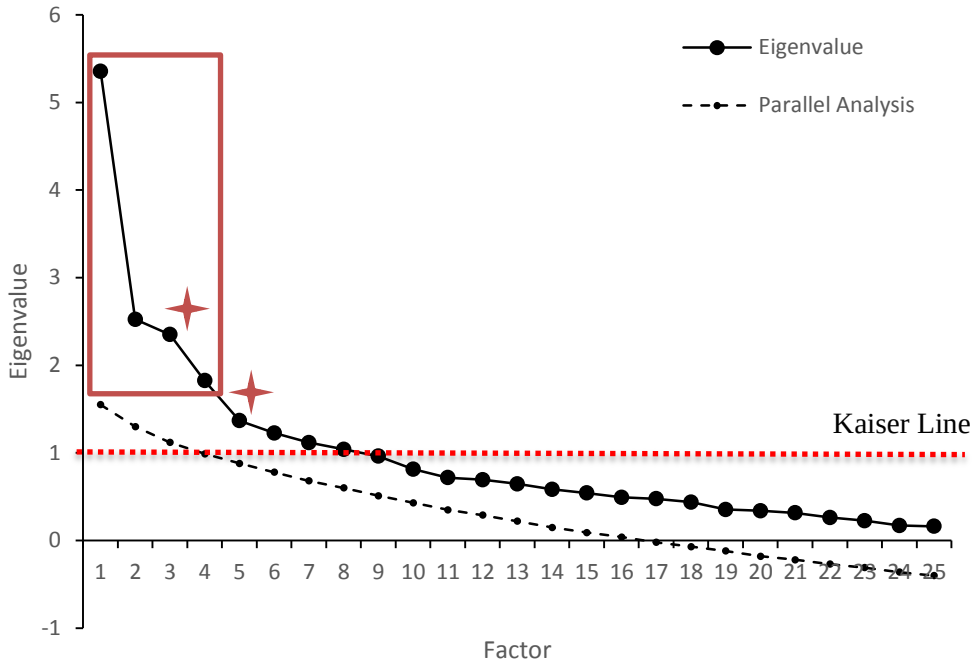


Figure 10: Scree plot and parallel analysis of eigenvalues for the parent informant data. ✦ indicates inflection points

Solutions for three, four and five factors were each examined using direct oblimin rotations of the factor loading matrix (Appendix 29). The four factor solution, which explained 48% of the total variance was preferred because of: (a) the 'levelling off' of the eigenvalues on the Scree plot after four factors (Figure 10); (b) some previous theoretical support (see section 2.5.1.4); (c) the difficulty of interpreting the five factor solution and (d) the apparent merging of the conceptually distinct Externalised (Conduct) and Internalised (Peer Problems) factors in the three factor solution

The Goodness-of-fit test ($\chi^2(206) = 253.49$, $p \leq 0.01$) indicates that there is more variance to be accounted for within the residual correlation matrix.

5.5.2.3 Factor Loadings for Parent Informant SDQ Data

Table 10 shows the rotated factor matrix of each item of the SDQ. The majority of the factor loadings of the individual items were highly significant.

A number (3) of items could be eliminated because they did not contribute to a simple factor structure and failed to meet a minimum criteria. The items 3 (Somatic), 4 (Shares) and 20 (Helpful) failed to load above 0.3 on any factor. Item 15 (Distractible) had a positive primary factor loading of 0.7 on Factor 1 and a positive cross-loading of 0.3 on Factor 3. Similarly, Item 2 (Restless) had a positive primary factor loading of 0.6 on factor 3 and a

positive cross-loading of 0.4 on factor 1. Item 1 (Considerate) had a negative factor loading of 0.3 on both factor 1 and factor 2.

5.5.2.4 Factor Correlations for Parent Informant SDQ Data

Although an oblique rotation was used, only small correlations between each of the factors existed none of which were significant (Table 10).

Table 10: Factor correlation matrix for parent informant data, based on a four-factor oblique rotation.

Factor	1	2	3	4
1		0.27	0.13	0.14
2			0.06	0.24
3				0.20
4				

5.5.2.5 Internal Consistency for Parent Informant SDQ Data

Internal consistency for each of the new factor composite scores was examined using Cronbach's alpha. Items with negative loadings were reversed coded prior to analysis. The Cronbach's alpha values were moderate: 0.76 for Hyperactivity (5 items), 0.67 for Conduct (5 items), 0.73 for Anxiety (5 items) and 0.70 for Peer Problems (7 items).

5.5.3 Factor Analysis of SDQ Self-report Data

5.5.3.1 Assumption Testing for Self-report SDQ Data

The factorability of the dataset was examined (see Section 4.7.4.4). All but one (Item 11. "Good Friend") of the 25 items correlated at least 0.3 with at least one other item. The overall Kaiser-Meyer-Olkin measure of sampling adequacy was 0.69. The majority of Kaiser-Meyer-Olkin values for the individual items exceeded the recommended value. The exception was for Item 11. (Good Friend; 0.37). Bartlett's test of sphericity was highly significant ($\chi^2(300)=734.93$; $p<0.001$). Five variables showed extraction communality values of less than 0.2 (Item 1 "Considerate", Item 4 "Shares", Item 11 "Good Friend", Item 14 "Popular", and Item 25 "Persistent") but were retained for comparability of analyses.

Overall, these preliminary analyses confirm the appropriateness of factor analysis with this data set.

Table 11: Factor loadings and communalities for the parent response data

Based on a maximum likelihood factor analysis with a direct oblimin rotation for 25 items from the strengths and difficulties questionnaire (SDQ; N = 101). Note: Correlations below 0.3 are not included. Red text indicate items that are loading on multiple factors. Dashes (-) indicate variables that fail to load onto any factor. (R) indicate this is a reverse worded item. Emboldened factor names represent a difference from the Goodman (1996) model and are discussed in section 2.5.1.4

	Factors				
	Self-regulation	Conduct problems	**Anxiety**	Peer problems	**Communality**
21. REFLECTIVE(R)_H	0.73				**0.55**
25. PERSISTENT(R)_H	0.72				**0.60**
15. DISTRACTABLE_H	0.70		0.35		**0.64**
7. OBEDIENT(R)_C	0.45				**0.51**
5. TEMPERS_C	0.34				**0.43**
12. FIGHTS_C		0.85			**0.62**
22. STEALS_C		0.72			**0.62**
18. LIES_C		0.57			**0.51**
9. CARING_PS		-0.38			**0.48**
1. CONSIDERATE_PS	-0.31	-0.38			**0.39**
20. HELPFUL_PS	-	-	-	-	**0.38**
10. FIDGETY_H			0.80		**0.71**
2. RESTLESS_H	0.45		0.61		**0.66**
8. WORRIES_E			0.54		**0.40**
24. FEARS_E			0.46		**0.46**
16. CLINGY			0.42		**0.31**
3. SOMATIC_E	-	-	-	-	**0.26**
11. GOOD FRIEND(R)_PP				0.61	**0.39**
14. POPULAR(R)_PP				0.58	**0.56**
23. BETTER WITH ADULTS_PP				0.55	**0.46**
19. BULLIED_PP				0.54	**0.39**
6. SOLITARY_PP				0.50	**0.49**
13. UNHAPPY_E				0.40	**0.50**
17. KIND_PS				-0.31	**0.39**
4. SHARES_PS	-	-	-	-	**0.36**
Eigenvalue	**5.36**	**2.53**	**2.35**	**1.83**	
% of total variance*	**21.43**	**10.10**	**9.40**	**7.30**	
Total variance				**48.22%**	

5.5.3.2 Number of Factors for Self-report SDQ Data

The scree plot (Figure 11) shows the number of factors with eigenvalues greater than 1.00. Following this rule would produce a 9 factor solution. The initial eigenvalues indicated that the first 3 factors explained 16%, 11% and 7% of the variance respectively. The fourth and fifth factors each explained between 5.09 – 5.93% of the variance. The sixth, seventh, eight and ninth factors each explained 4.16 – 4.95% of the variance.

The scree plot has two inflection points; one at factor 3 and a second at factor 5. The factors to the left of the inflection point were retained, i.e. the 3 largest factors.

A parallel analysis using a Monte Carlo simulation was used to determine the number of components or factors to retain when conducting the factor analysis. The analysis was run for the 122 eligible cases, with 25 variables, at a significance of 0.95, computing 1000 data sets, for common factor analysis using permutations of the raw data set. The matrix identified three significant components (Figure 11).

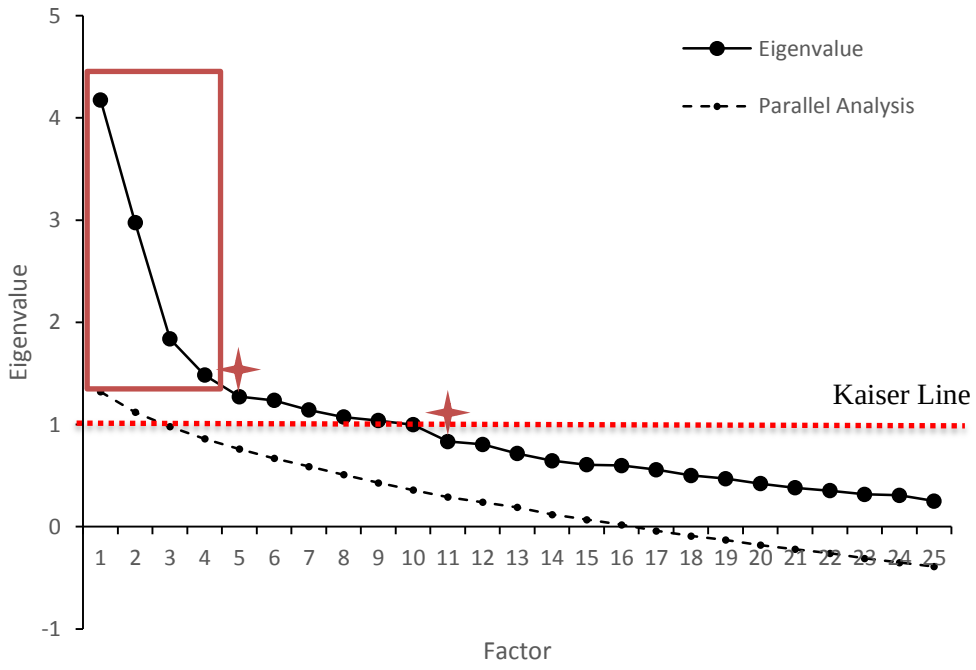


Figure 11: Scree plot and parallel analysis of eigenvalues for the self-report data. ✦ indicates inflection points

A factor solution with more than three factors was not possible, there were a number of "Heywood cases" (described in section 4.7.4.4). The Goodness-of-fit test ($\chi^2(228) = 273.914$, $p < 0.05$) indicates that there is more variance to be accounted for within the residual correlation matrix.

5.5.3.3 Factor Loadings for Self-report SDQ Data

Table 13 depicts the factor loading (after rotation) of each item of the SDQ. A number of items (3) could be eliminated because they did not contribute to a simple factor structure and failed to meet a minimum criteria. Three items (11. “Good Friend”, 20. “Helpful” and 14. “Popular”) failed to load above 0.3 on any factor. Item 9 (Caring) had a negative primary loading of 0.6 onto the “Extrovert” factor and a positive cross-loading of 0.3 onto the “Introvert” factor. Item 16 (Clingy) had a positive primary loading of 0.5 onto the “Introvert” factor and a negative cross-loading of 0.3 onto the “Extrovert” factor.

5.5.3.4 Factor Correlations for Self-report SDQ Data

Although a direct oblimin rotation was used, correlations between each of the factors were small and non-significant (Table 12).

Table 12: Factor correlation matrix for self-report data, based on a three-factor oblique rotation

Factor	Extrovert	Introvert	Hyperactivity
1		0.141	0.142
2			0.10
3			

Table 13: Factor loadings and communalities for the self-reported responses

based on a maximum likelihood factor analysis with a direct oblimin rotation for 25 items from the strengths and difficulties questionnaire (SDQ; N = 126). Correlations below 0.3 are not included. Red text indicates items that are loading on multiple factors. Dashes (-) indicate variables that fail to load onto any factor. Emboldened factor names represent a difference from the Goodman (1996) model and are discussed in section 2.5.1.4

	Factor			Communality
	Extrovert	**Introvert**	Hyperactivity	
9. CARING_PS	-0.63	0.39		**0.49**
12. FIGHTS_C	0.57			**0.45**
17. KIND_PS	-0.55			**0.34**
7. OBEDIENT(R)_C	0.54			**0.48**
22. STEALS_C	0.49			**0.46**
1. CONSIDERATE_PS	-0.44			**0.33**
18. LIES_C	0.41			**0.34**
21. REFLECTIVE(R)_H	0.35			**0.35**
4. SHARES_PS	-0.32			**0.34**
11. GOOD FRIEND(R)_PP	-	-	-	**0.25**
8. WORRIES_E		0.61		**0.42**
24. FEARS_E		0.60		**0.42**
16. CLINGY	-0.30	0.52		**0.37**
23. BETTER WITH ADULTS_PP		0.51		**0.32**
13. UNHAPPY_E		0.49		**0.38**
19. BULLIED_PP		0.45		**0.39**
5. TEMPERS_C		0.43		**0.37**
3. SOMATIC_E		0.42		**0.34**
6. SOLITARY_PP		0.42		**0.32**
20. HELPFUL_PS	-	-	-	**0.41**
14. POPULAR(R)_PP	-	-	-	**0.26**
2. RESTLESS_H			0.63	**0.34**
10. FIDGETY_H			0.51	**0.92**
15. DISTRACTABLE_H			0.46	**0.41**
25. PERSISTENT(R)_H			0.31	**0.34**
Eigenvalue	**4.18**	**2.97**	**1.84**	
% of total Variance	**16.70**	**11.90**	**7.35**	
Total Variance			**35.95**	

5.5.3.5 Internal Consistency for Self-report SDQ Data

Internal consistency for each of the new factor composite scores was examined using Cronbach's alpha. Items with negative loadings were reversed coded prior to analysis. The Cronbach's alphas were moderate: 0.67 for the "Extrovert" factor (9 items), 0.75 for the "Introvert" factor (9 items) and 0.60 for the Hyperactivity factor (4 items). No substantial increases in Cronbach's alpha for any of the scales could have been achieved by eliminating items.

5.5.3.6 Composite Scores and Re-analysis of the Self-report SDQ Data

The means (S.E) of the calculated factor scores for children, as rated by the child themselves are shown in Table 14. Composite scores were calculated using simple weighting (see section 4.7.4.4). Appropriate skewness and kurtosis indicate that these new composite scores are appropriate for parametric statistical analysis. When the results of the new factor scores were compared, there was a significant main effect of gender ($F(1, 117) = 3.80$, $p \leq 0.05$) for the 'Extroversion' sub-scale. The post hoc analysis revealed that girls rated themselves lower on the 'Extroversion' sub-scale compared to the boys (-1.04 ± 0.53, $p \leq 0.05$). There was no significant main effect of school year group, $F(1, 117) = 12.50$, $p = 0.22$) for the 'Extroversion' sub-scale. There were no significant interactions between gender x school year group for the 'Extroversion' sub-scale ($F(1,117) = 0.20$, $p = 0.66$).

Table 14: Descriptive data from the current sample using the three-factor structure

Factor	Gender	Mean (S.E)	Skewness	Kurtosis
Extrovert	Girls	-4.39 (0.33)	1.07	1.63
	Boys	-3.74 (0.41)	0.67	0.50
Introvert	Girls	6.86 (0.51)	0.63	-0.30
	Boys	6.62 (0.42)	0.30	-0.12
Hyperactivity	Girls	3.29 (0.25)	0.07	-0.72
	Boys	3.97 (0.23)	-0.26	-0.31

There were no main effects of gender $F(1, 117) = 7.41$, $p = 0.47$, school year group $F(1, 117) = 6.18$, $p = 0.51$ for the 'Introversion' sub-scale. There was no significant interaction between gender x school year group for the 'Introversion' sub-scale, $F(1, 117) = 3.11$, $p = 0.08$

There was no main effects of gender $F(1, 117) = 3.76$, $p = 0.06$, school year group $F(1, 117) = 0.58$, $p = 0.45$ for the 'Hyperactivity' sub-scale. There was no significant interaction between gender x school year group for the 'Hyperactivity' sub-scale, $F(1, 117) = 3.11$, $p = 0.08$ for the 'Hyperactivity' sub-scale.

Chapter Six: A Discussion of the Tool Development for this thesis

6.1 Overview of the Rationale for Tool Development

National and international policies are placing increased emphasis on the use of patient reported outcomes across all health services, including CAMHS. Although assessment has historically relied on adult informants, the drive for service user involvement has led to advances in the development of self-reported measures of wellbeing and behaviour in children and adolescents. However, the vast majority of research on self-reported measures of wellbeing and behaviour focuses on older children and adolescents. Nevertheless, these measures are often used clinically with younger children. Therefore, this study employed an observational design to compare the reports of behaviour and emotional difficulties of young children (8 to 10 years old) from multiple informants (child, parent and teacher) using a well-established measure (SDQ). The study aimed to establish the value of self-report measures in the accurate reflection of a child's behavioural and emotional difficulties. The study also aimed to assess the validity and reliability of the self-report measure, considering its potential use in a future nutritional intervention study. This was achieved by exploring the psychometric properties, reliability and validity of SDQ data collected from different informants; parents, teachers and the 8 to 10 year old schoolchildren themselves.

6.2 Summary of Results of Tool Development

The main findings are discussed below, with a focus on presenting the results within the context of the current literature, whilst considering the broader implications of (section 6.3) and learning gained (section 6.4) from this research.

6.2.1 Descriptive properties of the SDQ with a Community Sample of 8 to 10 year old Children

6.2.1.1 Effect of Informant

In this study the children reported more difficulties about themselves than either the parents and/or the teachers, across all sub-scales of the SDQ when using the 5-factor structure from Goodman et al. (1998). This is consistent with previous studies which have found that children and adolescents report more emotional and behavioural problems than their parents (Borg, Kaukonen, Joukamaa, & Tamminen, 2014; Capron et al., 2007; Johnson, Hollis,

Marlow, Simms, & Wolke, 2014; Mellor, 2005; Sawyer, Baghurst, & Mathias, 1992; Stanger & Lewis, 1993; Verhulst & Ende, 1992).

Given that this study assessed information from three informants, it is important to consider the potential for variation given the different investments from each party. The literature suggests that teachers are more likely to notice and report externalised difficulties (i.e. Hyperactivity and Conduct) for boys, in contrast to parents (Dobbs & Arnold, 2009). Similarly, teachers are more likely to notice prosocial behaviours in girls than in boys, in contrast to parents (Koskelainen et al., 2000; Koskelainen et al., 2001). Within the current dataset, parents reported more Internalising Difficulties (sum of Emotional Difficulties and Peer Problems sub-scales) and Externalising Difficulties (sum of Conduct Difficulties and Hyperactivity sub-scales), and Prosocial Strengths than did teachers. Differences in Internalising Difficulties is likely driven by differences in the Emotional Difficulties sub-scale, since parent and teacher reported Peer Problems did not differ. There was no evidence of an interaction between informant and gender for any of the sub-scales or composite scores.

6.2.1.2 Effect of Age

Within the current sample, the older children (Year 5) were reported to have more problems than the younger children (Year 4), by all informants. This appears inconsistent with previous findings which suggest that language deficits in younger children (aged 7 to 11) are associated with increased reports of behaviour problems (Norbury et al., 2016). Specifically, younger children are reported to have more hyperactivity problems than older ones (Woerner, Becker, & Rothenberger, 2004). Within this study it is important to note however, that both Year 4 and Year 5 groups were close in age and all are considered young children.

6.2.1.3 Effect of Gender

Boys, within this study, were reported to have more difficulties, independent of the informant type, than girls. This was consistent with previous studies that have found that parent reported Total Difficulty composite scores were higher for boys than for girls (Muris, Meesters, & van den Berg, 2003; Smedje et al., 1999; Woerner et al., 2004).

6.2.1.4 Endorsement Frequency of Likert Scale Responses

Likert-type scales are widely used among the general population because they offer an efficient method for capturing a wide range of response variance. However, completion of Likert-type scales is a complex task, requiring participants to distinguish between subtle differences in the frequency of attitudes or behaviours (e.g. ‘Some of the Time’ vs. ‘Often’ or ‘Always’). Little is known about whether young children can reliably and validly

categorize and distinguish subtle differences (e.g. ‘None’ to ‘A Little’, ‘Medium’ and ‘A Lot’) in their behaviours and attitudes on Likert-type scales.

Consideration of the endorsement rates for item level self-report responses within this study indicates that children have a tendency to choose the middle alternative “somewhat” over the extremes. This is consistent with findings from previous studies (Koskelainen et al., 2001) and is sometimes referred to as “piling on the midpoint” (Alreck & Settle, 1985; pg. 156). This can often occur when opinions are not firm or understanding is not clear, as might be the case for questionnaires beyond the developmental capabilities of young children (Stone et al., 2010). The consequence of this phenomenon is perhaps some attenuation of response reliability (Alwin & Krosnick, 1991; Weems & Onwuegbuzie, 2001). Another explanation may be associated with the level of attention difficulties identified within the sample. Inattention (specifically for the reverse word items) may have contributed to reduced understanding. However, the self-report endorsement rates for responses to reverse worded items were *not* consistently higher than the other informants.

The self-report data also shows that children are more likely to endorse “certainly true” responses for difficulties than either parent or teacher informants. In contrast, parent responses show a tendency to report “no problems”. This is consistent with findings from children aged 5 to 12 years (Chambers, Reid, Craig, McGrath, & Finley, 1998) and 5 to 11 years (Von Baeyer, Carlson, & Webb, 1997) which suggested that younger children have a tendency to endorse responses at the extreme end of scales when presented with items based on a Likert scale. Chambers and Johnston (2002) found that this was particularly evident when making judgements about feelings. This would fall in-line with classical theories of cognitive developmental (Piaget, 1954) which suggest that children in the stage of concrete operations (7 to 11 years of age; during which the child develops the capacity to make judgments and reason about the physical world) will have difficulties with these type of judgements, in contrast to children at the stage of formal operations (11 to 16 years; during which the children develops the capacity to think in abstract terms). The current findings demonstrate substantial differences in the way different informants respond to the questionnaire, and a concern might be that the use of a Likert-scale with young children may provide an unrefined measure of the constructs under investigation.

Unsurprisingly, for Likert-type data, the item responses were highly skewed. Some items were positively skewed (to the right), such as socially undesirable items i.e. Item 22 (Steals). While other items were negatively skewed (to the left), such as socially desirable items i.e. Item 11 (Good Friend) or Item 17 (Kind). The skewness is also likely to be a result of the relatively small sample size.

The items most likely to be endorsed as "certainly true" by teachers related to Hyperactivity Difficulties, with the exception of Item 10 (Fidgeting) and Item 8 (Worries; 11.7% - 18%; which related to Emotional Difficulties). Similarly, the items most likely to be endorsed as "certainly true" by parents were the five Hyperactivity items (14.9 – 25%). The children themselves were most likely to endorse difficulties associated with fear (Emotional Difficulties sub-scale), temper (Conduct Difficulties sub-scale), restlessness, fidgeting and distractibility (Hyperactivity sub-scale) as "certainly true" (22 – 30%).

6.2.2 Reliability of the SDQ with a Community Sample of 8 to 10 year old Children

The reliability of the SDQ within a community sample of 8 to 10 year olds, as reported by self-report, parent and teacher informant, was studied by exploring the internal consistency and inter-rater agreement correlations of the SDQ sub-scales.

6.2.2.1 Assessing Reliability of the SDQ using Inter-rater Consistency

The inter-rater agreement was assessed using two techniques. The use of correlation coefficients is perhaps the most commonly used method for calculating degree of consistency between informant type within the literature (see Appendix 31). However, this approach may be a poor reflection of the amount of agreement between informants, as it may result in extreme over or under-estimates of the true level of inter-rater agreement (McHugh, 2012; Stemler, 2004). Therefore, Cohen's kappa was also used, to account for the greater probability of agreement that arises when informants guess item responses.

6.2.2.1.1 Using Correlation Coefficients to assess Inter-rater Consistency

The correlation coefficients showed good agreement between parent and teacher informants, while agreement between the teachers and the children themselves was the lowest. This pattern is typical, based on the data from previous studies (Goodman, 2001; Goodman et al., 1998; Koskelainen et al., 2001).

The correlations for parent/teacher agreement were higher than those presented by Achenbach et al. (1987) in their meta-analytic study across a range of informants. They found the mean correlation for the parent/teacher reports was 0.27. Interestingly, within the current sample, the Conduct sub-scale produced a non-significant inter-rater agreement, whereas significant agreement between different informants has been observed on this scale in previous studies (Appendix 31). In contrast, the evidence seems to suggest that ratings of externalised problems lead to greater agreement between informants, relative to ratings of internalised problems (Stanger & Lewis, 1993).

For the parent/self-reports Achenbach et al. (1987) found a mean correlation of 0.25 and a mean correlation of 0.20 for teacher/self-reports, which is substantially higher than the agreement within this study. This is perhaps unsurprising within this younger sample, given that most studies looking at agreement between parent or teacher informants and the child themselves find a higher levels of agreement between the adult informants and the child as the age of the child increases (Niditch & Varela, 2011). The agreement between teacher informants and the children's own reports on the prosocial sub-scale was the highest. Interestingly, previous studies have found that agreement between informants on the Prosocial Strengths sub-scale, has been particularly low (Van Roy et al., 2008). It was suggested that this may represent difficulties with interpreting the meaning of the "prosocial behaviour". Palmieri and Smith (2007) have used these findings to suggest that the prosocial strengths sub-scale is a merely a methodological artefact, in that it functions as a positive construal factor (as it also contains loadings for positively worded items which should load onto other factors) .

6.2.2.1.2 Using Cohen's Kappa to assess Inter-rater Consistency

In contrast, the Cohen's kappa values suggest minimal inter-rater agreement between the informants across the sub-scales and composite scores. The agreement between parents and the children themselves when rating Emotional Difficulties was the only significant agreement. It is perhaps understandable that children may communicate their worries and sadness to their parents, rather than their teachers. However, the lack of agreement between informants for more objective and observable behaviours could indicate a discrepancy in the understanding of the constructs between the children, parents and teachers, guessing or differing responses to socially desirable responses by informants.

It is important to note that kappa values are very sensitive to skewed distributions, as is the case in the current data, so it may be that the low kappa values were to be expected (Vaz et al., 2016). Nevertheless, considered alongside the interclass correlations, the overall agreement across the five sub-scales between parent and teacher informants were appropriate, while the agreement between self-report and either parent or teacher informants was poor.

6.2.2.2 Assessing Reliability of the SDQ using Internal Consistency

The internal consistency of the SDQ Total Difficulty composite score, when informed by the teachers and parents, was satisfactory. However, when the SDQ Total Difficulty composite score was informed by the children themselves it showed much lower internal consistency. Importantly, Cronbach's alpha was higher than 0.70 for the Total Difficulty composite score which may be appropriate for this younger age group (Di Riso et al., 2010; Muris et al., 2004). It is recommended that scores yielding a minimum reliability of 0.7 - 0.8

(Ferketich, 1991; Lance et al., 2006) are acceptable. However, it is important to remember that these measures can be used as screening tools in the clinical context and as important decisions are being made based on these scores, perhaps a higher reliability of 0.9 might be advisable.

Across the sub-scales, the teacher informant data showed the highest average reliability. It may be that teachers are more reliable informants as they are an educationally homogeneous group and more reliably able to understand written questions, in contrast to parents who are from a range of social classes and educational backgrounds (Kresanov, Tuominen, Piha, & Almqvist, 1998). More specifically, the teachers may have training in dealing with a range of behaviours and the associated nomenclature which is used by the SDQ, whereas parents might vary in their understanding, perceptions and willingness to acknowledge that their child displays a behaviour they perceive negatively. This study did not collect data on the parents IQ or literacy level, as well as being a barrier to completion, this may have mediated the effects found. Interestingly, when reporting Hyperactivity, Total Difficulty and Internalised Difficulties, the parents were satisfactorily consistent informants. This pattern of high internal consistency for the Hyperactivity sub-scale is seen across studies (Goodman, 2001; Koskelainen et al., 2000; Malmberg, Rydell, & Smedje, 2003; Muris et al., 2003; Smedje et al., 1999; Woerner et al., 2004; see Appendix 32). The teacher informants within the current sample displayed greater internal consistency than that found in larger samples (Goodman, 2001; n=7313; Koskelainen et al., 2000; n=376). Additionally, the parent informants from the current sample produced internal consistency in line with previous studies (see summary in Appendix 32).

It is important to acknowledge that the use of the three-point Likert scale, as opposed to a five- or seven-point scale, will produce low levels of Cronbach's alpha (Zumbo, Gadermann, & Zeisser, 2007). Nevertheless, the children that completed the self-report versions of the SDQ within the current sample showed internal consistency similar to that found in comparable SDQ studies (Cronbach's alpha values range from 0.39 to 0.80; Goodman, 2001; Goodman et al., 1998; Koskelainen et al., 2001; Muris et al., 2003). The sub-scale relating to Peer Problems showed one of the lower consistencies, similar to self-report in previous studies (Palmieri & Smith, 2007).

6.2.2.3 Reliability Overall

In conclusion, the consistency of reports from different adult informants (parents and teacher) on the same measure of wellbeing and/or behaviour (SDQ) was consistent with previous research. The corresponding parent and teacher sub-scale scores correlated significantly. Furthermore, the internal consistency of the sub-scales, when informed by the adult informants was satisfactory and consistent with the current literature.

However, the self-reported sub-scale scores did not demonstrate the same level of agreement as the corresponding sub-scale scores produced by the adult informants. The internal consistency of the sub-scales, when informed by the children themselves was similar to the findings of comparable studies. Nevertheless, poor inter-rater agreement of the SDQ sub-scale scores between the children and their parents or teachers suggests that the self-reported SDQ sub-scales are not addressing the same constructs as those captured by the adult informants.

6.2.3 Validity of the SDQ with a Community Sample of 8 to 10 year old Children

The validity of the SDQ within a community sample of 8 to 10 year olds was assessed by studying the concurrent validity of the SDQ scales, as informed by the teacher, parent and the child themselves. Firstly, this was achieved by comparing internal construct sub-scale scores on the SDQ (such as Emotion and Peer Problems) with the corresponding sub-scale scores on the RCADS. Secondly, this was achieved by comparing behaviour sub-scale scores on the SDQ (such as Conduct and Hyperactivity) with the corresponding sub-scale scores on the SNAP-IV.

The composite scores and the sub-scale scores produced by the parent and the child themselves on the SDQ were compared to the sub-scale scores produced by the parent and the child themselves on the RCADS questionnaire (a measure of anxiety and depression; internalised difficulty). The correlations between the corresponding sub-scales on the SDQ and the RCADS, when informed by the parents were satisfactory. As would be expected, the correlations between the SDQ Emotional Difficulties sub-scale and the Total Difficulty composite score on the RCADS was the highest. Correlations between the SDQ Internal Difficulties composite score and the Depression sub-scale on the RCADS were also high. Correspondingly, the differences between the two questionnaires were highlighted by minimal correlation between the unique Prosocial Strength sub-scale on the SDQ and sub-scales on the RCADS. This suggests that the SDQ is likely to measure similar constructs to those measured by the RCADS in adults.

However, the sub-scales on the SDQ, when informed by the child themselves, did not significantly correlate with the sub-scales scores relating to anxiety and depression on the RCADS, despite being informed by the same child. This may indicate that the SDQ sub-scales are not representative of the same constructs measured using the RCADS (which is a measure of anxiety and depression symptomology established as valid and reliable in 8 to 10 year olds when rated by the child). Positively, the correlations between the SDQ Emotional Difficulties sub-scale and the Anxiety sub-scale on the RCADS showed some association, as did the correlations between (1) the SDQ Total Difficulty composite score

and the Depression sub-scale on the RCADS and (2) the SDQ Total Difficulty composite score and the Total Difficulty composite score on the RCADS. This may indicate a conceptual overlap between Total Difficulty and depressive symptomology and between emotional difficulties and anxiety.

The composite scores and the sub-scores produced by each informant (teacher, parent and the child themselves) on the SDQ were compared to the corresponding sub-scales produced by the parent and teacher informants on the SNAP-IV questionnaire (a measure of inattention and hyperactivity; ADHD symptomology and externalised difficulty). The correlations between the corresponding sub-scales on the SDQ and the SNAP-IV for the parent and the teacher data were satisfactory. The correlations between the SDQ Hyperactivity sub-scale and the SNAP-IV Inattention sub-scale were the highest, for both the parent and the teacher data. Similarly, the correlation between the SDQ External Difficulties composite score (an aggregation of Hyperactivity and Conduct sub-scales) and the SNAP Hyperactivity sub-scale were significant for both the parent and the teacher data. Although, it is important to note that the apparent overlap in construct of hyperactivity and conduct when informed by the adults, as this does suggest that for the adult informants the SDQ is likely to measure similar concepts to those measured by the SNAP-IV. Furthermore, there was no significant correlation found between the SDQ Peer Problem sub-scale and either sub-scale (Inattention or Hyperactivity) on the SNAP-IV when informed by the parents, which may support the discriminant validity of the measures. Conversely, the Hyperactivity and Inattention sub-scales on the SNAP-IV correlated with the internalised SDQ sub-scales, Emotional Difficulties and Prosocial Strengths.

When informed by the child themselves, none of the sub-scales of the SDQ significantly correlated with either of the sub-scales on the SNAP-IV when informed by the adults, perhaps implying a limited overlap of the ADHD contrasts between the child and the adults. Positively, the correlations between the self-reported SDQ Hyperactivity sub-scale and the Inattention sub-scale on the SNAP-IV completed by the parent informants did show the some association, perhaps suggesting the child's understanding of hyperactivity is more similar to their parents' concept of inattention. Interestingly, the association between Conduct and Hyperactivity seen with the adult informants was mirrored by the association between the self-reported SDQ Conduct Difficulties sub-scale and the Hyperactivity sub-scale on the SNAP-IV when informed by the parents. Unexpectedly, the teacher informed SNAP-IV Inattention and Hyperactivity sub-scales were mostly associated with the SDQ Prosocial Strengths sub-scale and the SDQ Emotional Difficulties sub-scale, respectively. This may suggest that teachers understand Prosocial and Emotional Difficulties as the reverse of Inattention and Hyperactivity.

In conclusion, the consistency of reports of the same construct by the adult informants, using different measures (SDQ, RCADS and SNAP-IV) was good. The corresponding SDQ Emotional Difficulties and Internal Difficulties sub-scale scores and the RCADS sub-scale scores correlated significantly when informed by the parent, as did the corresponding SDQ Hyperactivity and External difficulties sub-scale scores with the SNAP-IV sub-scale scores when informed by both the teacher and the parent. This suggests concurrent validity of the SDQ when informed by parent and teachers.

However, none the self-reported SDQ sub-scale scores correlated significantly with any sub-scale scores on the RCADS (Total, Anxiety or Depression) when informed by the parents or the children themselves. Nor did the self-reported SDQ sub-scale scores correlate significantly with any sub-scale scores on the SNAP-IV (Inattention or Hyperactivity) when informed by the parents or the teachers. This suggests that the SDQ when informed by the children themselves does not have sufficient concurrent validity. It is therefore likely that the SDQ when informed by the children themselves is measuring different facets of difficulty to the SNAP-IV and the RCADS when informed by the adults.

6.2.4 Exploratory Factor Analysis of the SDQ with a Community Sample of 8 to 10 year old Children

The original five factor structure of the SDQ as proposed by Goodman (1994) was assessed in a community sample of 8 to 10 year olds by exploratory factor analysis using data from self-report, as well as parent and teacher informants. This analysis explored the relationships between the questionnaire items to identify the constructs underlying the measure. The data collected from teachers was most consistent with the original factor structure outlined by Goodman (1994), followed by the parent reported data whilst the self-report data from the children themselves was least similar to the original factor structure of the measure.

6.2.4.1 Teacher Informant data

The validity of Goodman's (1994) original five factor structure of the SDQ was supported by the teacher data in the current study. The Goodness-of-fit test for the forced 5 factor model generated indicated some further variance may be accounted for within the residual correlation matrix. However, it is likely that these additional factors are weak, only explaining a very small amount of variance each and do not represent meaningful factors. It is important to note that, the goodness of fit statistic is also sensitive to sample size (Agrawal & Lord, 2006).

The factor correlation matrix produced (see Table 9) was consistent with the original sub-scale structure of the SDQ (Goodman, 1997). However, the Conduct Difficulties items load

negatively onto its factor. To explain this difference this factor has been renamed "propriety". This reversal is more representative of socially desirable behaviours, reflected in the re-naming of the factor. The exception to the predicted factor structure is Item 4. (Shares), which would typically be predicted to fall within the Prosocial sub-scale, but which has a primary positive loading on the newly named 'Propriety' sub-scale. The remaining Prosocial items load together and seem to signify socially desirable qualities or values, in contrast to behaviours.

The factor correlation matrix showed moderate positive correlations between factors. Correlations between the factors do not provide support for the discriminant validity of the factor constructs. This concern has been raised previously by Van Roy et al. (2008), who found high correlations between the Hyperactivity and Conduct sub-scales, and the Peer Problem and Emotional Difficulties sub-scale. This would be consistent with the three-factor solution suggested by Dickey and Blumberg (2004), who introduced the 'Externalised Difficulties' and 'Internalised Difficulties' sub-scales. They suggest that the five-factors are, at least in part, measuring the same underlying trait. Although the lowest correlation was found between the Prosocial Strengths and the Emotional Difficulties sub-scales, the significant correlation between the Prosocial Strengths sub-scale and the Hyperactivity sub-scale, does not wholly support the suggestion from Goodman et al., (2003) that "the absence of prosocial behaviour is conceptually different from the presence of psychological difficulties" pg. 174 (Goodman et al., 2003).

6.2.4.2 Parent Informant data

As seen previously with the teacher informant data, the Goodness-of-fit test for the forced 5 factor model indicated that there was remaining variance to be accounted for within the residual correlation matrix. However, it is likely that these factors are weak and do not represent meaningful factors. This was supported by the forced eight factor solution, which was not feasible due to a "Heywood case" (see section 4.7.4.4)

Within the current sample, the predicted five-factor structure did not fit the parent informant SDQ data. Instead, a four factor model provided a better fit (see Table 12). The current solution split the items previously thought to load onto the 'Conduct' sub-scale between two factors. The new 'Conduct' factor seemingly represents antisocial traits or behaviours, including fighting, stealing and lying (from the original Conduct sub-scale) in addition to negatively loaded items referring to socially desirable traits such as being caring and considerate of others (from the original Prosocial Strengths sub-scale). The new 'Hyperactivity' factor excludes fidgeting and restlessness (which now loads onto the Anxiety factor), but seemingly signifies aspects of self-regulation. This includes reflectiveness, persistence, distractibility (from the original Hyperactivity sub-scale),

obedience and temper control (from the original Conduct sub-scale). Interestingly, this four factor model might reflect the slight difference, between parents and teachers, in concurrent validity of the Hyperactivity sub-scale, which was seen when comparing the SDQ sub-scales with the SNAP-IV (see section 5.4). The factor labelled 'Anxiety' includes physiological symptoms such as fidgeting and restlessness (from the original Hyperactivity sub-scale) with psychological signs of anxiety such as worry, fear and nervousness (from the original Emotional Difficulties sub-scale). Dickey and Blumberg (2004) also found that when the factor rotation method allowed correlations between factors (as in the present study analysis) the external and internal dimensions merged. The final factor, 'Peer Problems', includes the five items from the original Peer Problems sub-scale, in addition to Item 12 (Unhappy; originally from Emotional Difficulties) and Item 17 (Kindness; originally from Prosocial Strengths). These additional items could be considered as social skills which facilitate the development of peer relationships.

Items referring to helpfulness and somatic symptoms failed to load onto any of these four factors. In a three factor solution, helpfulness loaded negatively alongside anti-social traits. Additionally, in a five factor solution, somatic symptoms loaded alongside emotional difficulties, such as worries and unhappiness. It is possible that parents find it difficult to identify the presence or absence of particular characteristics in their young children, making the endorsement of items difficult. Although not explored extensively within this study, this could potentially explain the missing data within the present dataset.

A four-factor model has been proposed as the best-fit model previously by Bull et al. (2016). They identified that the Hyperactivity item referring to fidgeting loaded onto Conduct Difficulties, while the Conduct item referring to obedience loaded negatively onto Prosocial Strengths. Emotional Difficulties and Peer Problems predominately loaded together onto an Internalising Difficulties scale, although peer items about sharing and helpfulness loaded onto the Prosocial Strengths sub-scale. It is interesting to note that these authors used the UK parent-informed preschool version of the SDQ (for a sample of children aged 52 to 87 months old) which removes antisocial conduct items about lying and stealing and replaces them with items indicative of being argumentative with adults and spiteful to others. Muris et al. (2004) also found support for a four factor model, although using self-report data from 8 to10 year old children, whereby Peer Problems and Conduct Difficulties items did not load onto separate factors.

It may also be interesting to note that the Bull et al. (2016) study used a multi-trait multi-method approach to modelling their data. The approach was developed by Campbell and Fiske (1959) to assess the concurrent validity of measures such as the SDQ. It highlights the distinction between convergent validity (where theoretical interrelatedness is evidenced) and

discriminate validity (where theoretical non-relatedness is evidenced). While to claim concurrent validity, both convergent and discriminate validity must be demonstrated. Within the context of the SDQ, items would therefore, load onto two factors: a trait factor (sub-scales) and a method factor (positive or negative; i.e. strengths or difficulties).

6.2.4.3 Children's Self-report Data

The predicted five-factor structure did not fit the self-report SDQ data in the current sample of children aged 8 to 10 years old. Instead a three factor model provided a better fit (see Table 14). In previous studies (Dickey & Blumberg, 2004; Koskelainen et al., 2001; Percy et al., 2008; Ruchkin et al., 2008), three factor solutions have distinguished between Internalised Difficulties (Emotional Difficulties and Peer Problems), Externalised Difficulties (Conduct Difficulties and Hyperactivity) and Prosocial Strengths.

In contrast, in the current model, Conduct items loaded positively alongside negatively loaded Prosocial items, indicative of anti-social behaviours. This factor also included one Hyperactivity item, relating to an ability to reflect on behaviour. The second factor was comprised of Emotional Difficulty items alongside Peer Problems, which mirrors the Internalised Difficulties and is perhaps indicative of items reflecting withdrawal or social isolation. Finally, the remaining Hyperactivity items are loaded together onto a factor specifically related to hyperactivity.

It is interesting that with younger children, the sub-scales are not distinguishing between strengths and difficulties. This is consistent with comments from Palmieri and Smith (2007), who highlighted that the prosocial sub-scale can function as a positive construal factor. It can vary based on raters' readiness to attribute positive qualities. This can also explain the high rate of positively worded (reverse coded) items cross-loading onto the original Prosocial factor.

It is proposed that the first two factors reflect binary self-categorisation as an extrovert or an introvert. In support, cognitive development literature suggest that younger children primarily think dichotomously (Gelman, Baillargeon, Mussen, Flavell, & Markman, 1983). Interestingly, previous research by Muris, Meesters, de Kanter, and Timmerman (2005) has found that child-reported measures of behavioural inhibition (conceptually similar to introversion), from 8 to 12 year olds, appeared significantly connected to self-report of the SDQ Emotional Difficulties sub-scale, whereas child-reported measures of behavioural activation (conceptually similar to extroversion) were related to self-report of the SDQ Hyperactivity/Conduct problems. In another study Muris, Meesters, and Diederen (2005) found associations between extraversion and the SDQ prosocial sub-scale, in children 12 to 17 years old. This has also been confirmed in a sample of Russian adolescents aged 10 to 18 years old (Slobodskaya, 2007).

An important aspect of this dichotomy is the role of social desirability. A common problem in determining the validity of measures is that individuals try to present themselves in the most favourable light. Children (and their parents) are possibly more likely to provide biased subjective opinions, and show a tendency to avoid showing themselves, or their family, in a negative light. The prevalence of socially desirable responses in self-report data increases the risk of spurious correlations between items, as well as the suppression or moderation of the relationships between constructs. This effect could have been emphasized by the nature of the data collection. The context of the school setting may have emphasized the children's desire to provide the "correct" responses, heightened by concerns of discussing personal problems with a stranger, the presence (and sometimes open communication) of peers, the gender divide (both boys and girls were supported by myself, a female researcher) and the norms of adult (and academic) authoritarianism over the younger children. However, the level of difficulties reported for the children were highest when reported by children themselves, which suggests that self-report SDQ results are not subject to social desirability and demand characteristic biases.

Importantly, the new SDQ composite scores derived from the factor analysis show appropriate psychometric characteristics underlining its suitability for future use with community samples of young children, aged between 8 and 10 years old.

6.2.4.4 Overall Findings from the Exploratory Factor Analysis of the SDQ, when informed by Teachers, Parents and the Child themselves

The structural validity of the teacher informant data supported the original five factor structure of the SDQ as proposed by Goodman (1994). The parent informant data supported a four factor structure. However, this did not appear to be a result of Internalising factors (Emotional Difficulties and Peer Problems) merging onto one factor, unlike the four factor solutions previously identified in the literature (Bull 2016).

The self-reported data produced a three factor solution, which appears to reflect a dichotomy in young children's categorisation of behaviours and traits of: (1) active, extroversion-like and (2) inactive, introversion-like behaviour, alongside (3) a factor considering hyperactivity. This three-factor structure produced consistent composite scores appropriate for future analysis.

6.3 Strengths and Clinical, Theoretical and Research Implications of the Developing the Tools for this Thesis

The findings from this thesis can be considered important for a number of reasons. Firstly, this thesis is novel in its examination of the psychometric properties, the reliability and

validity of the SDQ, when used with multiple informants, to assess a UK sample of healthy children aged 8 to 10 years.

Within clinical practice, the SDQ is currently used with children under the age of 11. Hence, the findings of this study suggest that the SDQ self-report version has different factors than those currently considered when it is completed by younger children (aged 8 to 10 years old). This means that the current sub-scales may be less appropriate or reliable when completed by this younger group. Therefore, it is important to carefully consider the psychometric properties of outcome measures that are being used in child psychiatry and psychology.

6.3.1 Clinical Implications of the Tool Development

These findings have a number of clinical implications for the interpretation of data provided by the self-report version of the SDQ when used with young children (8 to 10 years old) and, also, more generally, for the understanding of the use of self-report measures of behaviour and wellbeing. Importantly, when employing the self-report SDQ in children below age 11 years, clinicians and researchers should at least ensure that the child comprehends the questionnaire items and the rating scale. It is important not to overlook the risk associated with pathologising misunderstood reports or even potentially transient developmental behaviours in young children. The use of clinical judgement will therefore, remain a vital part of the exploration of emotional and behavioural difficulties in children.

The findings have serious implications for the use of self-report measures with young people in clinical practice. The SDQ is useful as a general screening instrument in community settings (e.g. schools) to serve as a baseline of a children's emotional and behavioural state, when used alongside the adult informant versions. However, the new three-factor structure suggested by this study could improve the reliability of the self-report version. If this measure were available for use in this age range it could facilitate the development of the evidence base for interventions with young children, through its potential use as a routine outcome measure.

The incremental validity of multiple informants has recently been considered by Aitken, Martinussen, and Tannock (2017), who found that ratings collected across settings by multiple informants can add useful information, but they emphasised that it may not be necessary or an effective use of resources to use self-report or parent informed ratings when teacher symptom ratings are available. Rønning, Handegaard, Sourander, and Mørch (2004) recommended that a minimum of two informant versions should be used. They warn against using the self-report version in isolation, as self-report questionnaires are subject to various biases (such as social desirability). It has long been established that information from multiple informants is a better predictor of disorder than information relying on just one

source (Meltzer, Gatward, Goodman, & Ford, 2000; 19-20pg), as it should also be noted that adult informant are also subject to social desirability bias. For example, Sanne, Torsheim, Heiervang, and Stormark (2009) found substantial differences in the response from teachers and parents to SDQ Item's 4 (Shares) and 11 (Good Friend) in an attempt to appear socially desirable. Similarly, Janssens and Deboutte (2009) found significant correlations between SDQ sub-scale scores and parent/caregiver demand for additional support, which suggests that this measure is influenced by the informant's motivation to achieve a particular outcome. The impact of social desirability has been discussed in cross-cultural studies using a translated version of the SDQ (Mansbach-Kleinfeld et al., 2010), considering the impact of culture on the importance of social perception. Socially desirable could have be amplified by the nature of a self-selected sample, it is possible that parents willing to engage in research are also prone to socially desirable responding.

6.3.2 Theoretical Implications of the Tool Development

Additionally, it is likely that the variability in the number of factors generated by data from different informants is linked to the multidimensional nature of the behaviours being identified. For example, ADHD is a behavioural disorder characterised by symptoms such as inattention, hyperactivity and impulsiveness (APA, 2013). Broken down, the SDQ only includes two items on inattention, two items on hyperactivity and one items on impulsiveness. The limited number of items for each construct is likely to contribute to poor internal consistency and contradictory factor structures. Similarly, the lack of uni-dimensionality of the sub-scales means that two or more factors are contributing to each scale, resulting in unpredictable scale behaviour (Mellor & Stokes, 2007). For example, Item 16 (Clingy), which was originally included on the Emotional Difficulties sub-scale, refers to anxiety is social situations, which is likely to impact upon or be closely associated with Peer Problems. With this in mind, it is important for clinicians to use their clinical judgement and follow up the SDQ with disorder specific measures.

6.3.3 Research Implications of the Tool Development

The results of this study provide additional information on the usefulness of the SDQ self-report version as a tool for epidemiological research. Although it is recommended that the findings are replicated. The findings suggest that the psychometric properties of measures of behaviour and wellbeing vary across age-bands and informant type. Hence, this is likely to have implications for the outcome measures selected (and the scoring of those measures) to assess the efficacy of interventions designed to improve behaviour and wellbeing in young children.

6.4 Limitations of the Research Developing the Tools for this Thesis

There were several limitations that should be considered when interpreting these findings.~~,~~ For example, with regards to factor analysis, Tabachnick and Fidell (2007) suggest having at least 300 cases (pg. 613) but that a sample size of 150 cases is sufficient if solutions have several high loading marker variables (>0.80). However, with factor loadings between 0.31 – 0.87, only a maximum of 126 cases were available for the analysis. Furthermore, in developing the tools, the examination of the concurrent validity of the SDQ was done by comparing parent/teacher responses to RCADS and SNAP-IV. However, the parent and teacher response rates were lower than the children's, which may explain some of the inconsistencies in the findings.

Furthermore, no age comparison group was used in this research. Results were collected from only two year groups, while it might have been beneficial to compare the younger sample to an age group considered to be valid and reliable self-reporters (e.g. aged 11 to 16 years old) according to the established evidence base (Goodman et al., 2003). Additionally, no clinical comparison group was used in this research, which would have allowed for the assessment of the SDQ self-report measure's ability to differentiate clinical cases in a younger sample. Thus providing information pertaining to the predictive validity of the self-report SDQ. The probable clinical cases (that is, those children and adolescents who require further assessment and, possible treatment or other interventions) were likely to be identified as outliers, defined from scoring distributions on the SDQ scales. Additionally, it would have been possible to gather information about common psychosocial problems not covered using the SDQ or other contributing factors (e.g family structure) to consider the impact of these elements on the children's understanding of the constructs measured by the questionnaire.

There were also some methodological concerns around the use of Cronbach's alpha when examining the reliability of measures which are ranked on a Likert scale (e.g. the SDQ) have been raised, with mean inter-item correlation coefficients based on polychoric correlations suggested instead (D'Souza, Waldie, Peterson, Underwood, & Morton, 2017). Likert scale scores are normally considered to be ordinal or discrete data (rather than continuous; Jamieson, 2004; Norman, 2010). Polychoric correlation is a technique for estimating the correlation between two theorised normally distributed continuous latent variables, from two observed or ordinal variables (Kolenikov & Angeles, 2009). It is also know that Cronbach's alpha values are affected by scale length and therefore interpretation of individual sub-scales may be low because of the small number of items (5) which make up the five SDQ sub-scales (D'Souza et al., 2017; Streiner, 2003). This was not accounted for.

A key aspect of this research was that each child's difficulties were rated by multiple informants (parents, teachers and the children themselves). These scores were then compared when calculating inter-rater reliability. Typically inter-rater reliability is calculated from observations of the same material by different informants, for example when two physiatrists are independently diagnosing disorders based upon the same videotaped consultation. In contrast, in the present study, each informant had access to only part of the whole information about the child, within certain contexts. For example, the child knows his or her feelings and ideas better than others would, while parents see how their child behaves at home and with family members, while teacher see how the child behaves at school. Goodman (1997) suggested that given this, intra-class correlations should not be used for reliability estimates. So perhaps the reliability of the SDQ data reported in the studies within this thesis should be treated with caution.

6.4.1 Future Research Directions informed by the Development of the Tools for this Thesis

The use of the SDQ self-report, specifically its factor structure, might be considered in other populations. The prevalence of emotional and behavioural difficulties in children with intellectual disability is much higher than that of their typically developing peers, while there is comparatively little information available on the use of self-report measures by young people with intellectual disabilities. Currently the SDQ has not been validated within this population. Goodman, Ford, Richards, Gatward, and Meltzer (2000) advised that the SDQ (self and informant report versions) is not appropriate for use with children experiencing severe/profound disabilities, as self-injurious and autistic features are not captured (Law & Wolpert, 2014). However, Haynes et al. (2013) found support for a three factor structure (dissimilar from the current findings and comprising Positive Relationships, Negative Behaviour and Emotional Competence) of the self-report version, while Emerson (2005) reported it to be a robust measure of behaviour and wellbeing in this population.

6.5 Overall Summary and Conclusions Following the Tool Development

6.5.1 Psychometric Properties of the SDQ in a Community Sample of Young Children

Examining the structural validity of the SDQ when completed by young children (8 to 10 years old) is critical. These measures are being routinely used in CAMHS services across the country with little consideration for the variable validity and reliably across type and age of informant (e.g. parents vs. teachers vs. adolescence vs. young children). This chapter has

presented information concerning the SDQ when informed by three informants (teachers, parents and the children themselves) in a sample of young (8 to 10 years old) schoolchildren.

In conclusion, the internal reliability, inter-rater agreement and concurrent validity of the SDQ for measuring the behaviour and wellbeing of young children (8 to 10 years old), when completed by different adult informants (parents and teacher) was satisfactory and consistent with previous research. The structural validity of the teacher informant data supported the original five factor structure of the SDQ as proposed by Goodman (1994), while the parent informant data supported a four factor structure, suggesting some discrepancies in construct perception between raters. The current findings suggest that overall, perhaps, the teacher data is the most useful.

While the internal consistency of the sub-scales, when informed by the children themselves was appropriate and similar to the findings of comparable studies, the self-reported sub-scales did not demonstrate sufficient agreement with corresponding sub-scale scores produced by the adult informants. Similarly, the self-reported SDQ sub-scale scores did not correlate significantly with any sub-scale scores on the RCADS when informed by the parents or the children themselves. It was suggested that the SDQ when informed by the children themselves is measuring different facets of difficulty, as compared to when informed by adults. This was in part confirmed by the assessment of the structural validity of the SDQ. The self-reported data produced a three factor solution, which appears to reflect a dichotomy in young children's categorisation of behaviours and traits of: (1) active, extroversion-like and (2) inactive, introversion-like behaviour, alongside (3) a factor considering hyperactivity. This three-factor structure produced internally consistent composite scores appropriate for further analysis.

This suggests that, while the self-report version of the SDQ was designed for children and adolescents aged 11 years and over, consistent with the data from Di Riso et al. (2010) the measure can provide valid information about behaviour and wellbeing in children as young as 8 years old, provided that the method of scoring the questionnaire is adapted.

6.4.2 Using the SDQ in Intervention Research

Alongside aims of developing evidence-based treatment interventions for clinical difficulties in behaviour and wellbeing, there is also a need to develop methods to assess the efficacy of these interventions. Robert Goodman, the developer of the SDQ, is a member of the CAMHS Outcome Research Consortium (CORC), which collaborates with CAMHS across the U.K. Their aim is to encourage a common model of routine outcome evaluation. The current CORC handbook (www. corc.uk.net) recommends the SDQ as one of the routinely used methods for evaluation of interventions in child and adolescent psychiatric

clinical practice. It is therefore, relevant to assess whether the SDQ is appropriate for intervention evaluation in younger samples. This thesis, therefore, used the tool developments outlined above to evaluate the effects of a dietary intervention on the behaviour and wellbeing of healthy, 8 to 10 year old children using the SDQ alongside other measures of behaviour and wellbeing. Correspondingly, this thesis explored the effects of a dietary intervention using the revised factor structure for the self-reported data. The three-factor structure was analysed in a sample of 8 to 10 year olds to ensure improved reliability over the original five-factor model in this younger population.

Chapter Seven: Thesis Aims - Exploring the Effect of Flavanone-rich Dietary Intervention on Behaviour and Wellbeing of Healthy, Young Children using the SDQ

The main aim of this thesis was to explore the effects of 10 weeks daily consumption of flavanone-rich orange juice compared with placebo (orange flavoured drink) on behaviour and wellbeing measures in 8 to 10 year old schoolchildren.

7.1 Hypotheses

1. 10 weeks of daily consumption of a flavanone-rich orange juice drink, compared to a placebo drink, will lead to a change in:
    a) Reports of behaviour and wellbeing difficulties, as represented by the SDQ Total Difficulty sub-scale score, as reported by parents, teachers or the child themselves.
    b) Reports of strengths, as represented by the SDQ Prosocial Strengths sub-scale score, as reported by the parents, teachers or the child themselves
    c) Reports of anxiety and depression, as represented by the total RCADs score, as reported by parents or the child themselves
    d) Reports of behaviours associated with ADHD, as represented by the SNAP-IV sub-scale scores, as reported by parents and teachers.
2. Following the development of appropriate tools for use in this thesis, the chronic (10 week) consumption of a flavanone-rich orange juice drink, compared to a placebo drink, will lead to a change in reports of behaviour and wellbeing, as represented by the three newly developed sub-scales (Extroversion, Introversion or Hyperactivity) of the SDQ self-report version.

Chapter Eight: Methodology

8.1 Design

This study employed a randomised, chronic, 10 week, parallel groups, double blind, placebo controlled design to investigate the effects of daily consumption of a flavanone-rich 100% orange juice drink compared with an equicaloric placebo drink (orange flavoured drink) on behaviour and wellbeing in 8 to 10 year old schoolchildren.

8.2 Ethical Clearance

The study received ethical approval from the School of Psychology Research Ethics Committee (SoPREC) at the University of Leeds. The SoPREC reference numbers and dates of approval are: 16-0022 approved 19/01/2016 (main study, with a focus on cognitive measures – data reported elsewhere), 16-0181 approved 29/06/2016 (addition of the behavioural and wellbeing measures, an optional sub-study and the focus of this thesis), 16-0187 approved 16/07/2016 (addition of the gut microbiota measures, an optional sub-study – data reported elsewhere) and 16-0349 approved 28/11/16 (amendment to permit a second cohort).

8.2.1 Intellectual property

The RCT was funded by the State of Florida, Department of Citrus (FDOC), and the University of Leeds retained the right to use data generated in the course of conducting the study for research, education, patient care purposes and publishing results of the study consistent with the terms of contracted services and without payment of royalties or other fees.

8.3 Participants

8.3.1 Recruitment

An opportunity sample of children aged 8 to 10 years were recruited to participate in this study. The children were recruited from two primary schools in the local area, from a pool of children from two academic years. Ages 8 to 10 years correspond to compulsory primary school years 4 and 5 in the British school system. Both years correspond to Key Stage 2 in the British education system.

This age range was chosen in regards to cognitive performance and responsiveness to nutritional intervention. Studying this age group permitted analysis on a focussed and narrow age range. It represents an important time period for children. Academic achievement during primary school years is fundamental and the focus is on working towards Key stage 2 national curriculum examinations at age 11, at the end of primary school education. In addition, it has been suggested that a peak in brain growth occurs at the age of 7 years (Epstein, 1986) and the period of 7 to 9 years is critical for development of executive functioning (Anderson, 2002). Nutritional interventions are likely to have greatest effects on brain function during periods of growth and development (Keunen, Van Elburg, Van Bel, & Benders, 2014). This age range represents a time when nutritional interventions may be particularly effective as the central nervous system may be most sensitive to nutrient benefit when it is developing

Appendix 18 shows the flow diagram of recruitment. This study was powered for the primary cognitive outcome variables of the main study. It was not possible to power the study on the behaviour and wellbeing measures as there is no existing data on a dietary intervention using the selected measures.

The study was powered on picture recognition based on the studies of nutritional interventions in children of a similar age below. The estimated sample size required for this study was based on the following assumptions: (a) 5% significance level, (b) 80% power, and (c) equal number of participants in each intervention or treatment group. This was informed by the data reported by Taib et al (2012), in which an intervention of three different types of milk with glucose or isomaltulose were administered to 5 to 6 year old children. In comparison to baseline, it was estimated that the number of participants required per group to detect an effect of the magnitude reported in Taib et al (2012) was less than 15 for most of the outcome measures they employed. To detect a change from baseline on the numerical working memory speed outcome in the isomaltulose enriched growing up milk (iso GUM) group and the glucose group, the required sample size was estimated to be 20. It was also estimated that the required sample size to detect the effect reported by Taib et al (2012) when comparing the difference between intervention groups, was less than 20 per group to detect significant differences between the standard GUM vs. iso GUM, and reformed GUM vs. iso GUM, with the exception of picture recognition sensitivity between standard GUM vs. iso GUM which requires 32 participants per group.

Furthermore, using data from a snack intervention study by Muthayya et al (2007) in 7 to 9 year old children, it was also estimated that the required sample size per group for this study (based on the same assumptions as above) was 26 participants per group to detect the difference (5-15%) between the treatment group relative to the control on immediate picture

recognition with a standard error of 1.5%. Muthayya et al (2007) based their sample size calculation on the study of Busch et al (2002), which also confirmed the sample size calculation of the current study, in that the required sample size to detect a mean difference of 4.3 with standard deviation of 4.8 was 30 participants per group.

Therefore, a target number of 100 participants were to be recruited in order to complete with 40 in each arm of the study (allowing for a dropout of 20% in each arm). This was thought to exceed the required sample size and allow for detection of effects on each cognitive outcome variable without computing composite cognitive function scores

For the analysis of gut microbiota data, it was estimated that a sample size based on a study by Pawitan et al (2005) would be appropriate. This study prescribed the estimation based on the false discovery rate (FDR) control. The false discovery rate is the expected proportion of 'false positives' among the list of genes that we consider significant. The current study used the following assumptions: (a) 50% of the microbiomic profile is not associated with the case-control status of the orange juice intervention and (b) at least 80% power. The calculation indicated that with a sample size of 20 in each experimental group, it was expected to control FDR between 6% and 13%. At sample size 30 in each experimental group, it was expected to control FDR between 0.5% and 4%. FDR of 13% is not considered high in an experiment as discussed by Pawitan et al (2005) as it can go up to more than 50%. Hence, it was considered that a sample size of 20 in each group would be adequate to control FDR at reasonable levels. Therefore, a smaller sample was considered appropriate for the sub-studies (including the current thesis).

8.3.2 Consent

Before consent was sought, the pupils were told about the research study via an assembly held by the researchers at the schools. Following this assembly, parents/carers received the letter and information sheet sent to their home to inform them about the study (Appendices 19-28). This detailed the requirement for participants to consume the orange juice drink at weekends and during the school holidays.

Parent and teacher information sessions were offered at each school for parents/carers/teachers to attend should they have any questions or queries – although this was not taken up. The teaching staff were given detailed explanations of what the study entailed and how their classes would be affected from the head teachers in school meetings.

8.3.2.1 Cohort One

For the first cohort, consent for the main RCT (with a focus on cognition) was obtained for 56 children. Consent for the additional (1) behaviour and wellbeing and (2) gut microbiota sub-studies was gained from parents/carers in the form of opt-in permission letters sent

home with the children and consent was obtained from each individual child on each day of testing.

An honorarium was given for completion of the main RCT (including the behaviour and wellbeing measures) in the form of shopping vouchers (£20.00).

8.3.2.2 Cohort Two

For the second cohort, consent for the full study was gained from parents/carers in the form of one opt-in permission letter).

An honorarium for the completion of the full study in the form of shopping vouchers (£50.00) was offered. Furthermore, class teachers who were asked to complete the measures were offered an honorarium of shopping vouchers (£2.00 per complete set of questionnaires; two per child at pre- and post-intervention)

8.3.3 Participant Screening

Potential participants attended a screening session prior to the scheduled baseline testing and administration of the intervention. During this, the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1991) was administered to measure Intelligence Quotient (IQ). IQ was therefore included as a covariate in the analysis of all outcomes.

Each participant had to meet all of the following inclusion criteria and none of the exclusion criteria to participate in the study. Inclusion criteria: Girls or boys, aged 8 to 10 years old, ability to follow verbal and simple written instructions in English, ability to give written informed consent and ability to understand cognitive written instructions and measures. Exclusion criteria: Learning disabilities which interfere with the ability to understand written or verbal communications, visual impairment that precludes the ability to follow read written instructions and measures, Inability to understand the objective of the measures, or complete the items, acute illness, or feelings of unwell, within the week prior to testing and current administration of any psychotropic, antibiotic medication or supplementation in the month prior to testing, or during testing and any food allergies or intolerances.

At the end of the study, participants were asked to rate how much they liked the orange juice drink that they received during the study using a child friendly 10 point Likert scale.

8.3.4 Randomisation

Following screening, participants were randomised into the active or placebo orange flavoured drink group. The randomisation procedure was developed by Dr Gusnanto (the study statistician). Randomisation lists were created for participants within each school class and according to gender to achieve the most balanced assignment to study treatments.

Consideration was given to the randomisation of siblings within the study (i.e. siblings received the same treatment to avoid contamination of between active and placebo treatment arms). Researchers remained blind to the allocation of the active and placebo drinks until completion of the data analysis. Participants, parents and teachers were also blind to their assigned condition.

8.4 Intervention

Following screening, participants were randomly assigned to one of two conditions: (1) Active group: orange juice, (2) Placebo group: orange drink. Both the orange juice and placebo drinks were provided by the sponsor in blind packaging. The cartons were labelled as either Red (Active) or Blue (Placebo). Table 15 provides the nutritional composition of the active and placebo juice drinks. On average, an 8oz serving of either juice drink contained around 100-120 kcals.

Table 15: Nutritional composition of the orange juices (mg / 8 oz)

ppm = parts per million, mg/8oz = milligrams per 8oz serving

	Active		**Placebo**	
Brix[7] (Corrected)	11.91		11.91	
Brix : Acid Ratio[8]	16.89		17.94	
	Ppm	**mg/8oz**	**ppm**	**mg/8oz**
Glucose (mg)	30211	7159.9	31047.9	7358.34
Fructose (mg)	32533.6	7710.45	36161	8570.15
Sucrose (mg)	57078.6	13527.63	47157.6	11176.36
Vitamin C (mg)	468	100.92	450	106.65
Hesperidin (mg)	426	100.96	0	0
Narirutin (mg)	70.7	16.76	0	0

Drinks were stored at the school in fridges and handed out to the children at school or in cool bags for weekend and holiday consumption (which they could retain at the end of the study). The children were informed that the drinks should be transferred to home fridges for consumption on that day or weekend to ensure the drink was chilled and palatable

8.4.1 Outcome Measures

The effects of 100% Florida Orange Juice on measures of children's behaviour and wellbeing were examined using the self-report, parent and teacher versions of the Strengths

[7] Brix is a unit used to describe the percentage of soluble solids in juice (primarily sugars and some acids), this is corrected for acidity and temperature.

[8] The ratio of the Brix value of the juice to the grams of citric acid of the juice.

and Difficulties Questionnaire (SDQ; Goodman, 1997; see section 2.5.1), the self-report and parent versions of the Child Anxiety and Depression Scale (RCADS; Chorpita et al., 2005; Muris et al., 2002; see section 2.5.2) and the parent and teacher versions of the 18-item SNAP-IV (Bussing et al., 2008; see section 2.5.3), a revised version of the Swanson, Nolan and Pelham (SNAP) Questionnaire measuring ADHD. Data were collected at baseline and following a chronic (10 week) intervention with the orange juice supplement.

8.5 Procedure

8.5.1 Study Visit Schedule

The study schedule ran over 11 weeks; one week for screening/baseline measures, plus a 10 week intervention period, followed by a two week follow-up period for the return of questionnaires (see Figure 12). Data were collected from two cohorts in two data collection phases (Cohort 1: September to December 2016 and Cohort 2: January to March 2017).

8.5.2 Setting

All screening, the self-report measures and the administration of the test products was conducted in the school environment alongside the normal school day. As previously discussed, conducting the study within the school environment aimed to emulate the participant's usual routine as much as possible and to permit researchers to monitor compliance to the intervention. Moreover, this design aligned with previous school food policy schemes which parents were familiar with, such as the Free School Milk Scheme provided in primary education until 1996; current milk schemes in primary education that provide free or subsidised milk for children and the School Fruit and Vegetable Scheme which entitles children aged 4-6 years to receive a free portion of fruit or vegetables each day.

However, the procedure also required participants to consume juice on non-school days during the weekend period and during school holidays at the usual time (mid-morning).

The informant-rated measures were completed outside of school hours, in either online (Bristol Online Surveys) or paper format

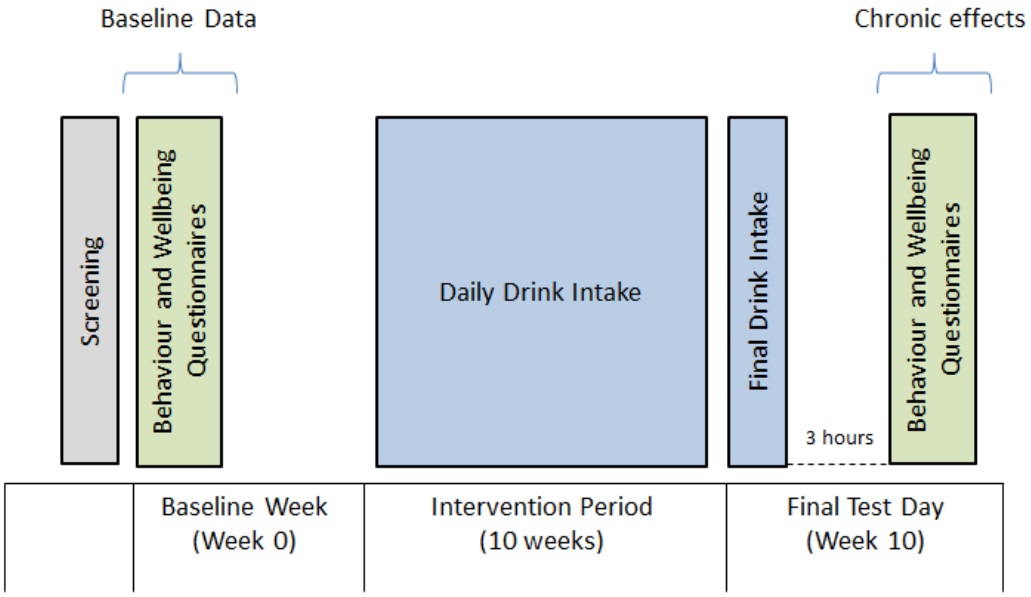


Figure 12: Study schedule schematic

8.5.3 Administration of the Treatment

Both groups received a chronic, 10 week, daily, supplemental drink of orange juice or placebo. These were delivered as a school-based mid-morning intervention administered by researchers. At this time, the researchers observed the children to monitor intake, compliance and to ensure there was no sharing, swapping or spillage of drinks.

Spare cartons of juice were provided to children for planned absences from school. Advice was provided to the children and their parents to consume the drinks mid-morning if they were not at school. Similarly, during non-school days (weekends) and during school holidays, cartons were supplied to the children on Fridays at school in cool bags for consumption at home. A drink diary log sheet was included, on which parents or children could tick off each day's drink or note non-compliance. This was removed and replaced by the study staff when the drink bags were replenished.

8.5.4 Statistical Analysis

For details on the treatment/processing of the questionnaire data, see section 4.7.2. The data were analysed separately for each informant (teacher, parent and the child).

8.5.4.1 Mixed Methods Analysis

The current data has characteristics of longitudinal data; it has both between-subjects and within-subjects variation, time-dependent covariates and missing data (Davis, 2002). Observations from the same subject are likely to be correlated, which violates the

assumption of independent observations of parametric statistical methods (Diggle, 2002). A linear mixed-effects model can accommodate these features (Tabachnick & Fidell, 2007).

A linear mixed-effects model was used to model the sub-scale score trajectory over time, with timepoint (pre / post) considered as the time variable. The continuous outcome, sub-scale scores, were measured at the start and end of the intervention. Therefore, the sub-scale scores for each participant were expected to be correlated and the linear mixed-effects model is most appropriate for taking into account both within-subject variation and between-subject variation. In the present study, age and IQ are continuous covariates included in the models. Gender, treatment and time point are categorical variables. A restricted maximum likelihood (REML) method for parameter estimation was used (Tabachnick & Fidell, 2007).

The procedure for model selection used a backwards step-wise process. All main effects and interactions were entered in the first model and the model fit, F values and significance of main effects and interactions examined. Non-significant interactions were removed, starting with highest order interactions, and the resulting model was compared to the previous model using the McQuarrie Tsai AICc criterion (McQuarrie & Tsai, 1998). The AICc criterion gives an indication of the amount of remaining unexplained variance after the model has been fitted, where a smaller AICc value indicates a better model fit. This was used in preference to the Akaike's information criterion (AIC) because the AICc protects against overfitting. If an improvement in model fit was found, other non-significant effects were removed and again the AICc criterion was used to evaluate the model fit. Models were chosen on the basis of 'best fit', and interaction terms that improved the fit were retained. Where Homogeneity of variance is not assumed, Greenhouse Geisser values were used. Whole numbers were reported for degrees of freedom and annotated with "gg adj" as subscript in each case. Main effects were explored using the Bonferroni corrected comparisons of marginal means. Effects of covariates were explored through consideration of scatterplots which plotted the covariate against the outcome variable, separately for treatment or other fixed factor effects where appropriate. In the event of a significant interaction, the interaction was plotted on the scatterplot with the variable and sub-scale score, with best fit regression lines plotted for each intervention group. The direction and strength of the relationship (positive or negative) is also be described.

8.5.4.2 Factor Scores

Composite scores for the SDQ self-report data were created in line with the newly developed factor structure (see sections 5.5.3 and 6.2.4.3). The calculation method is described in section 4.7.4.5.

Chapter Nine: Results

This chapter presents the data analysis and results for the intervention study. The effect of the intervention on measures of behaviour and wellbeing, was assessed using linear mixed-effects modelling (LME; see section 8.5.4.1). The purpose is to fit a model to estimate the effects of treatment, gender, age and IQ on the mean value and age change trajectory for the dependent variables of interest (SDQ Total Difficulty composite score and Prosocial Strengths sub-scale score). The final section assesses the effects of the intervention on the SDQ self –report measure, using the factor structure outlined in Chapter Four.

9.1 Participant Characteristics

Data were collected from 32 children at two time points (pre and post 10 weeks of dietary intervention; n = 66 total; see Table 17). Of these children, 44 SDQ's were completed by parents and 50 were completed by teachers across two schools at two time points. The mean age (S.E) of the sample was 111.26 months (6.67) which is equivalent to 9 years and 3.2 months. The age data were not normally distributed (Shapiro-Wilk test = $P<0.01$). A Mann Whitney U test revealed that the ages of the treatment groups were significantly different ($U=55.50$, $p < 0.05$). Table 16 shows that the Active group was on average 5.84 months younger than the Placebo group. IQ data were normally distributed. Independent t-tests revealed that there were no significant differences between the groups in terms of IQ ($t(29)=0.58$, $p > 0.05$).

Table 16: Mean (±S.E) age, IQ by treatment group

	Placebo Group		**Active Group**	
	N	**Mean (S.E)**	**N[9]**	**Mean (S.E)**
Age (months)	14	114.43 (1.36)	17	108.59 (1.67)
IQ	14	97.79 (2.70)	17	100.65 (3.86)

[9] Participant data for age was missing for one participant

		Active									Placebo								
		Self-report			**Parent**			**Teacher**			**Self-report**			**Parent**			**Teacher**		
		Pre	**Post**	**Change**	**Pre**	**Post**	**Change**	**Pre**	**Post**	**Change**	**Pre**	**Post**	**Change**	**Pre**	**Post**	**Change**	**Pre**	**Post**	**Change**
SDQ	*N*	18	18		12	10		17	14		14	14		11	11		12	7	
Total Difficulty	***Mean***	9.44	8.56	**-0.88**	6.92	6.70	**0.22**	5.06	4.93	**-0.13**	12.42	13.07	**0.65**	11.36	10.73	**-0.63**	8.08	7.57	**0.51**
	S.E	1.14	1.00		1.74	1.62		1.34	1.38		1.67	1.33		1.74	1.90		1.84	2.83	
Prosocial		8.33	8.28	**-0.05**	8.92	9.00	**-0.08**	7.59	8.14	**-0.55**	7.43	7.21	**-0.22**	8.18	8.18	**0.00**	6.75	6.29	**0.46**
		0.45	0.47		0.31	0.56		0.55	0.49		0.54	0.60		0.44	0.50		0.75	0.68	
Emotion		**2.89**	**2.22**	**-0.67**	**2.08**	**1.90**	**-0.18**	**1.65**	**1.50**	**-0.15**	**2.79**	**3.00**	**0.21**	**3.27**	**2.00**	**-1.27**	**1.50**	**3.14**	**1.64**
		0.40	0.49		0.51	0.71		0.59	0.51		0.65	0.53		0.82	0.63		0.29	1.22	
Conduct		**1.61**	**1.28**	**-0.33**	**1.50**	**0.70**	**-0.80**	**0.35**	**0.36**	**0.00**	**2.64**	**2.86**	**0.21**	**1.91**	**2.18**	**0.27**	**1.67**	**1.57**	**-0.10**
		0.36	0.29		0.48	0.30		0.15	0.17		0.49	0.49		0.44	0.42		0.67	0.87	
Hyperactivity		**2.94**	**2.78**	**-0.17**	**2.58**	**2.80**	**0.22**	**1.94**	**1.79**	**-0.16**	**4.00**	**4.57**	**0.57**	**4.27**	**4.27**	**0.00**	**3.67**	**2.86**	**-0.81**
		0.44	0.40		0.74	0.70		0.52	0.49		0.63	0.58		0.73	0.62		0.89	1.26	
Peer Problems		**2.00**	**2.28**	**0.28**	**0.75**	**1.30**	**0.55**	**1.12**	**1.29**	**0.17**	**3.00**	**2.64**	**-0.36**	**1.91**	**2.27**	**0.36**	**1.25**	**0.57**	**-0.68**
		0.41	0.36		0.35	0.52		0.34	0.64		0.65	0.48		0.53	0.69		0.39	0.30	
RCADS		18	18		12	10		-	-	-	14	14		11	11		-	-	-
Total Depression		**7**	**4.89**	**-2.11**	**2.67**	**2.10**	**-0.57**	**-**	**-**	**-**	**8.71**	**8.64**	**-0.07**	**3.82**	**2.82**	**-1.00**	**-**	**-**	**-**
		1.38	1.20		0.69	0.91		-	-	-	1.42	1.49		1.21	0.99		-	-	-
Total Anxiety		**9.33**	**7.61**	**-1.72**	-	-	-	-	-	-	**11.00**	**9.43**	**-1.57**	-	-	-	-	-	-
		1.70	1.99		-	-	-	-	-	-	1.51	1.90		-	-	-	-	-	-
Total All		**16.33**	**12.50**	**-3.83**	**16.75**	**12.30**	**-5.45**	-	-	-	**19.71**	**18.07**	**-1.64**	**21.73**	**18.09**	**-3.64**	-	-	-
		2.94	3.08		3.81	2.91		-	-	-	2.60	3.11		6.04	4.87		-	-	-
SNAP-IV	*N*	-	-		9	9		16	14		-	-		11	9		11	7	
Inattention		**-**	**-**	**-**	**0.8**	**0.48**	**-0.32**	**0.84**	**0.59**	**-0.25**	**-**	**-**	**-**	**1.04**	**0.86**	**-0.18**	**0.77**	**0.57**	**-0.2**
		-	-	-	0.27	0.13		0.25	0.18		-	-	-	0.26	0.23		0.24	0.27	
Hyperactivity		**-**	**-**	**-**	**0.47**	**0.29**	**-0.18**	**0.23**	**0.25**	**0.02**	**-**	**-**	**-**	**0.72**	**0.68**	**-0.04**	**1**	**0.62**	**-0.38**
		-	-	-	0.19	0.1		0.17	0.13		-	-	-	0.19	0.24		0.34	0.32	
Combined		**-**	**-**	**-**	**0.63**	**0.39**	**-0.24**	**0.68**	**0.42**	**-0.26**	**-**	**-**	**-**	**0.88**	**0.77**	**-0.11**	**0.88**	**0.6**	**-0.28**
		-	-	-	0.22	0.11		0.19	0.12		-	-	-	0.18	0.2		0.28	0.29	

Table 17: Mean (± S.E) of the sub-scale scores as rated by each informant, by treatment group and time point. Plus the change in these from pre to post intervention

Interaction	**Total Difficulty** T	**Prosocial** T	**Total Difficulty** P	**Prosocial** P	**Total Difficulty** SR	**Prosocial** SR	**Extraversion** SR	**Introversion** SR	**Hyperactivity** SR
treatment x timepoint	X	X	X	X	X	X	X	X	X
timepoint x gender	X	X	X	X	X	X	X	X	X
timepoint x age	X	X	X	X	X	X	X	X	X
timepoint x IQ	X	X	X	X	X	X	X	X	X
treatment x gender	X	X	X	X	X	X	X	X	X
treatment x age	X	X	$p < 0.001$	X	X	X	X	X	X
treatment x IQ	X	$p < 0.01$	X	$p < 0.01$	X	$p < 0.05$	$p < 0.01$	X	$P < 0.01$
treatment x timepoint x gender	X	X	X	X	X	X	X	X	X
Main Effects									
Treatment	X	X	$p < 0.01$	$p < 0.01$	X	X	X	X	X
Timepoint	X	X	X	X	X	X	X	X	X
Gender	$p < 0.05$	$p < 0.01$	$p < 0.01$	X	$p < 0.01$	X	$p < 0.01$	X	$p < 0.01$
Covariate Main Effects									
IQ	X	X	X	X	X	X	X	X	X
Age	$p < 0.01$	X	$p < 0.05$	$p < 0.05$	X	X	X	X	X

Table 18: Summary table of significant main effects, covariate effects and interactions for the SDQ.

Key: T = Teacher, P = Parent, SR = Self-report. R indicates this was removed from the final model (not applicable); X indicates that this was included in the final model, but was non-significant; * p = 0.05, ** p < 0.01. Reported effects were only present with the outliers removed

9.2 SDQ – The Effect of Dietary Intervention

Data were analysed separately by type of informant (teacher, parent and the child themselves). Table 18 summarises these findings.

9.2.1 SDQ Total Difficulty Sub-Scale

9.2.1.1 SDQ Total Difficulty Sub-Scale Teacher Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Total Difficulty composite score when informed by teachers revealed that there were no main effects of timepoint or treatment (see Appendix 34). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that teachers scored girls has having lower Total Difficulty scores on the SDQ compared to boys (mean difference -4.78 ± 1.83, $p < 0.05$). IQ was not a significant covariate. Age was a significant covariate such that age was positively related to the SDQ Total Difficulty composite score with the score increasing as age increased (Figure 13). There were no significant 2-way or 3-way interactions, apart from the interaction.

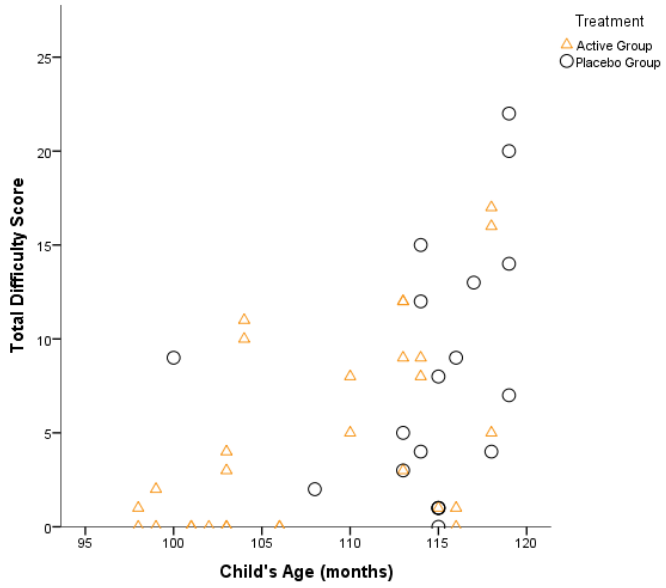


Figure 13: Scatterplot depicting child's age against SDQ Total Difficulty composite score as informed by the teacher, labelled by treatment group

9.2.1.2 SDQ Total Difficulty Sub-Scale Parent Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Total Difficulty score

when informed by parents revealed that there was no main effect of timepoint (see Appendix 34). However, there was a significant main effect of treatment. Bonferroni corrected post hoc analysis showed that the parents of children receiving the Active drink rated their children as having lower Total Difficulty on the SDQ compared to parents of children receiving the Placebo drink (mean difference -6.96 ± 1.58, $p < 0.001$). There was also significant main effect of gender. Bonferroni corrected post hoc analysis showed that the parents of girls rated their children as having lower Total Difficulty on the SDQ compared to parents of boys (mean difference -4.81 ± 1.54, $p < 0.001$). IQ was not a significant covariate. Age was a significant covariate such that age was negatively related to the SDQ Total Difficulty with the score decreasing as age increased (Figure 14). There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x age[1] (Figure 14). Younger children receiving the Placebo juice were rated as having more difficulties compared to those in the Active group, but there is no difference in Total Difficulty score for older or younger children the Active group.

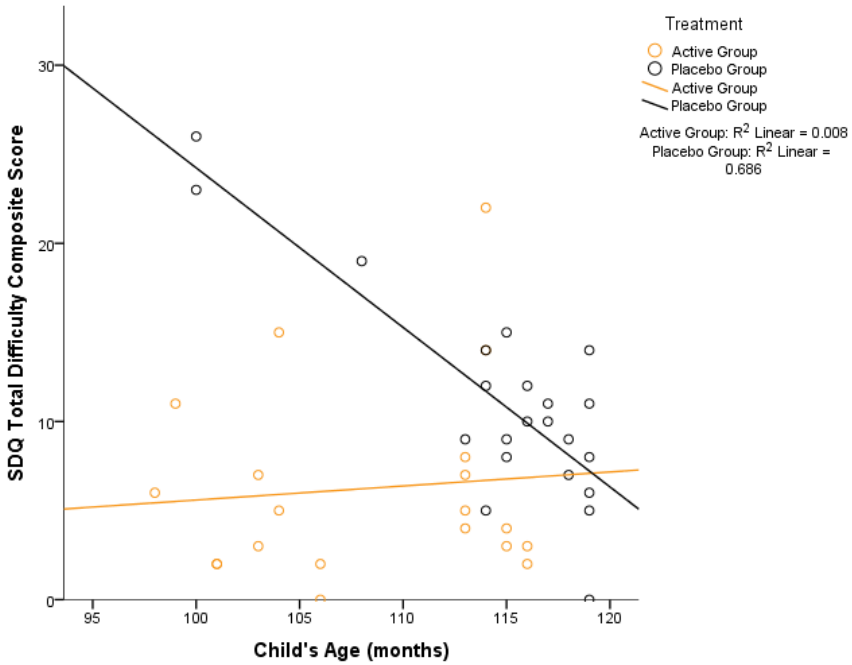


Figure 14: Scatterplot depicting child's age against SDQ Total Difficulty score as informed by the parent, labelled by treatment group. Linear fit line for treatment groups is shown

9.2.1.3 SDQ Total Difficulty Sub-Scale – Children's Self-Report

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Total Difficulty

[1] The removal of outliers did not change the analysis outcome

composite score when completed by the child revealed that there were no main effects of timepoint or treatment (see Appendix 34). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that girls rated themsleves as having lower Total Difficulty scores on the SDQ compared boys (mean difference -4.43 ± 1.43, $p < 0.001$). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions.

9.2.2 SDQ Prosocial Strengths Sub-scale

9.2.2.1 SDQ Prosocial Strengths Sub-scale Teacher Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Prosocial Strengths sub-scale when informed by teachers revealed that there was no main effect of timepoint or treatment (see Appendix 34). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that the teachers rated girls as having higher Prosocial Strengths on the SDQ compared to boys (mean difference 1.64 ± 0.62, $p \leq 0.01$). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x IQ (Figure 15). Children within the Placebo group with higher IQ scores were rated by their teacher as having lower Prosocial Strengths, compared to those in the Active group. There was no difference in scores on the Prosocial sub-scale for children with low IQ scores, irrespective of treatment group.

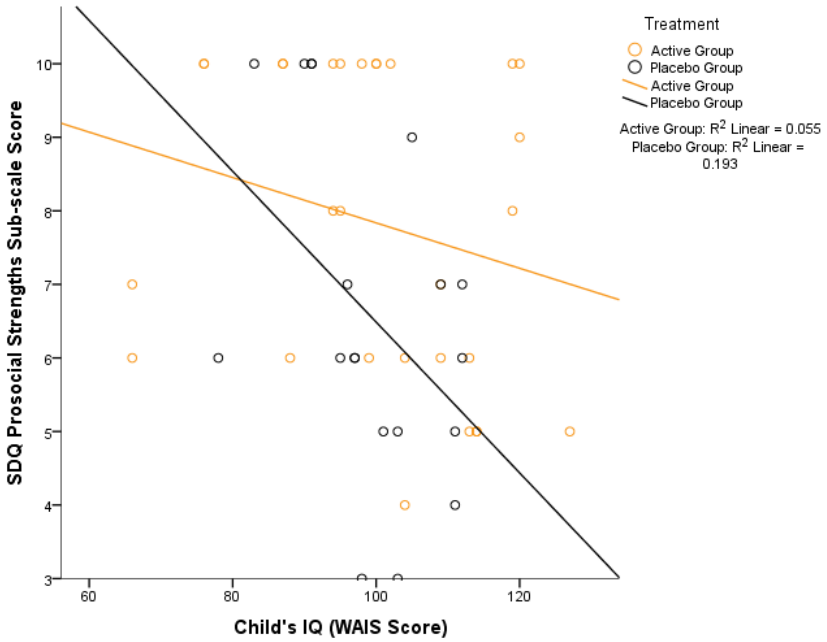


Figure 15: Scatterplot depicting child's IQ (represented by WAIS score) against SDQ Total Difficulty score as informed by the teacher, labelled by treatment group. Linear fit line for treatment groups is shown

9.2.2.2 SDQ Prosocial Strengths Sub-scale Parent Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Prosocial Strengths when informed by parents revealed that there was no main effect of timepoint or gender (see Appendix 34). However, there was a significant main effect of treatment. Bonferroni corrected post hoc analysis showed that the parents of children receiving the Active drink rated their children as having higher Prosocial Strengths on the SDQ compared to parents of children receiving the Placebo drink (mean difference 1.30 ± 0.48, $p \leq 0.01$). IQ was not a significant covariate. Age was a significant covariate such that age was positively related to the SDQ Prosocial Strengths sub-scale with the score increasing as age increased (Figure 16).

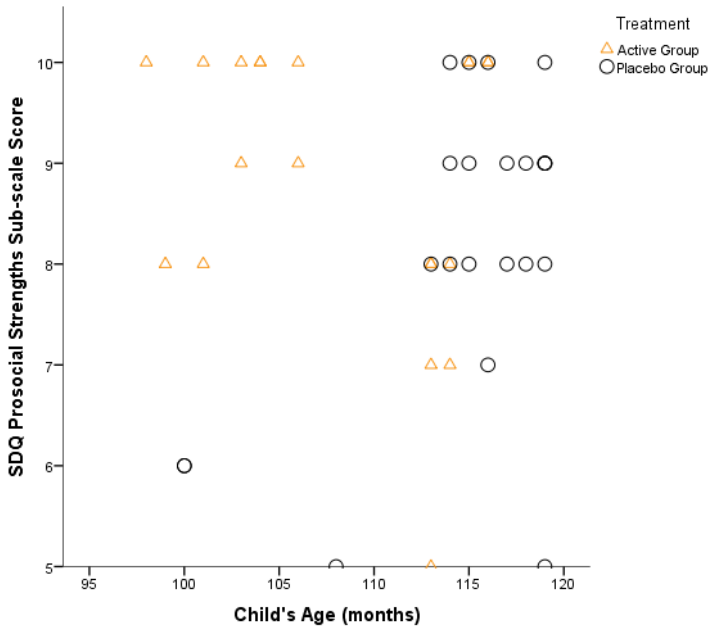


Figure 16: Scatterplot depicting child's age against SDQ Prosocial Strengths sub-scale score as informed by the parent, labelled by treatment group

There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x IQ (Figure 17). Children in the Placebo group with lower IQ scores were rated by their parents as having lower Prosocial Strength scores, while children in the Active group with lower IQ scores were rated by their parents as having more Prosocial Strengths.

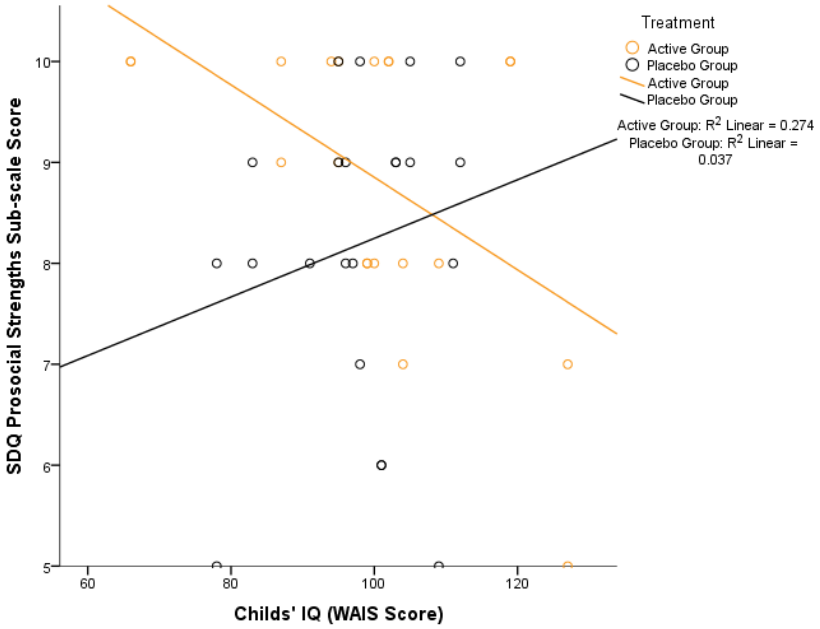


Figure 17: Scatterplot depicting child's IQ (represented by WAIS score) against SDQ Prosocial sub-scale score as informed by the parent, labelled by treatment group. Linear fit line for the treatment groups is shown.

9.2.2.3 SDQ Prosocial Strengths Sub-scale Children's Self-Report

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Prosocial Strengths when completed by the child revealed that there was no main effect of timepoint, treatment or gender (see Appendix 34). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x IQ (Figure 18). Children in the Placebo group with lower IQ scores rated themselves as having higher Prosocial Strength scores, while there were no differences the ratings of Prosocial Strength for the children in the Active group.

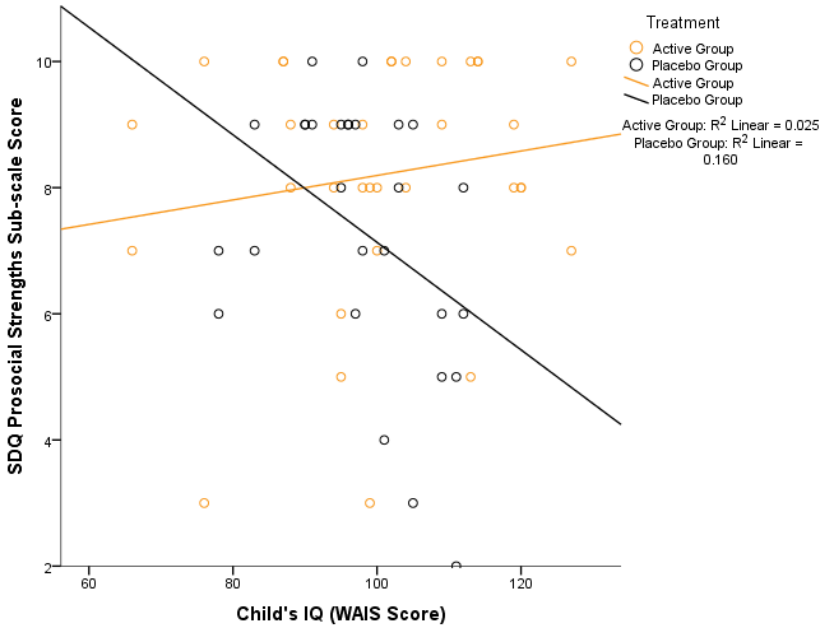


Figure 18: Scatterplot depicting child's IQ (represented by WAIS score) against SDQ Prosocial sub-scale score as informed by the children themselves, labelled by treatment group. Linear fit line for treatment groups is shown.

9.2.3 SDQ Self-Report with Young Children – Three-Factor Model Sub-scales

When developing the tools for this thesis, it was found that the five-factor structure of the SDQ was not appropriate for the use in younger children and a three-factor structure was derived, composite scores based on these three factors were also analysed to examine the effect of the 10 week dietary intervention.

9.2.3.1 SDQ Self-Report with Young Children – Extrovert Sub-scale

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Extrovert sub-scale (discussed in Section 6.2.4.3) when informed by the child revealed that there was no main effect of timepoint or treatment (see Appendix 35). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that girls rated themselves as having lower Extrovert scores on the SDQ compared boys (mean difference -2.17 ± 0.68, $p \leq 0.01$). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x IQ (Figure 19). Children in the Placebo group with higher IQ scores rated themselves higher on Extroversion, than children in the Active group with similar IQs. There was no difference in scores on the Extroversion sub-scale for children in the Active group.

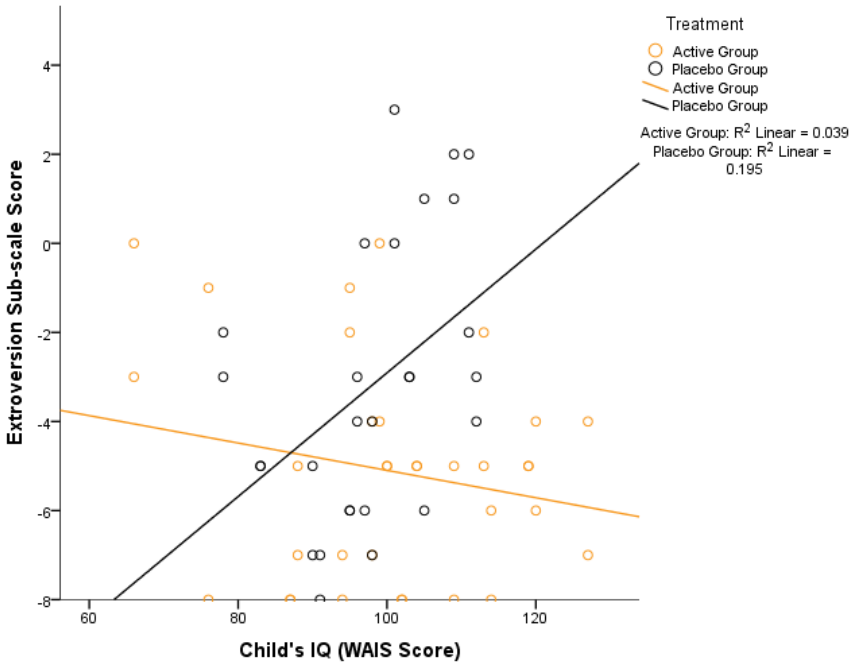


Figure 19: Scatterplot depicting child's IQ (represented by WAIS score) against SDQ Extroversion sub-scale score (based upon the newly developed three factor solution) as informed by the children themselves, labelled by treatment group. Linear fit line for treatment groups is shown.

9.2.3.2 SDQ Self-Report with Young Children – Introvert Sub-scale

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Introvert sub-scale (discussed in Section 6.2.4.3) when informed by the child revealed that there was no main effect of timepoint, treatment or gender (see Appendix 35). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions.

9.2.3.3 SDQ Self-Report with Young Children - Hyperactivity Sub-scale

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SDQ Hyperactivity sub-scale (discussed in Section 6.2.4.3) when informed by the child revealed that there was no main effect of timepoint or treatment (see Appendix 35). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that the girls rated themselves as having lower Hyperactivity scores on the SDQ compared boys (mean difference -1.67 ± 0.55, $p \leq 0.01$). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x IQ (Figure 20). Children in the Placebo group with higher IQ scores rated themselves higher on Hyperactivity compared to children in the Active group with similar

IQs. There was little difference in scores on the Hyperactivity sub-scale for children in the Active group.

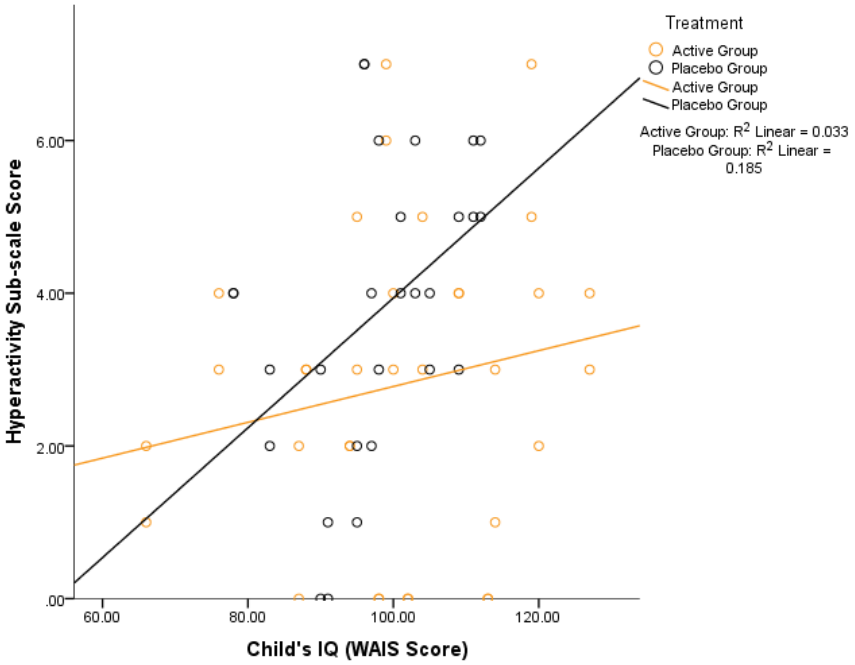


Figure 20: Scatterplot depicting child's IQ (represented by WAIS score) against SDQ Hyperactivity sub-scale score (based upon the newly developed three factor solution) as informed by the children themselves, labelled by treatment group. Linear fit line for treatment groups is shown.

	RCADS		**SNAP-IV**			
	RCADS Total Score		**Inattention**	**Hyperactivity**	**Inattention**	**Hyperactivity**
Interaction	P	SR	T	T	P	P
treatment x timepoint	X	X	X	X	X	X
timepoint x gender	X	X	X	X	X	X
timepoint x age	X	X	X	X	X	X
timepoint x IQ	X	X	X	X	X	X
treatment x gender	X	X	X	$p < 0.05$	X	X
treatment x age	X	X	X	X	X	X
treatment x IQ	X	X	X	X	X	X
treatment x timepoint x gender	X	X	X	X	X	X
Main Effects						
Treatment	X	X	X	X	X	X
Timepoint	X	X	X	X	X	X
Gender	X	X	$p < 0.001$	$p < 0.001$	$p < 0.01$	X
Covariate Main Effects						
IQ	X	X	X	X	X	X
Age	X	X	$P < 0.001$	$P < 0.001$	X	X

Table 19: Summary table of significant main effects, covariate effects and interactions for the RCADS and SNAP-IV.

Key: T = Teacher, P = Parent, SR = Self-report. R indicates this was removed from the final model (not applicable); X indicates that this was included in the final model, but was non-significant; * $p = 0.05$, ** $p < 0.01$. Reported effects were only present with the outliers removed

9.3 RCADS – The Effect of Dietary Intervention

Data were analysed separately, by informant (parent and the children themselves). Table 19 summarises these findings.

9.3.1 RCADS Total Score

9.3.1.1 RCADS Total Score Parent Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the RCADS Total score when informed by parents revealed that there was no main effect of timepoint or gender (see Appendix 36). However, there was a significant main effect of treatment. Bonferroni corrected post hoc analysis showed that the parents of children receiving the Active drink rated their children as having lower Total scores on the RCADS compared to parents of children receiving the Placebo drink (mean difference-12.23 ± 4.75, $p < 0.05$). IQ was not a significant covariate. Age was a significant covariate such that age was negatively related to the RCADS Total score with the Total RCADS score decreasing as age increased. There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x age which seemed to be driven by parent reported total RCADS scores for one child. The values were not a data entry error and represented values within the range of the measure (Pre – 78, Post – 50). Hence, it is thought that the values may represent an child with clinically relevant anxiety levels (a raw score of >75 in Girls aged 8 to 10 years old, means that the score is roughly in the top 2% of scores of un-referred young people of the same age; Chorpita et al., 2005). Nevertheless, the parent reported post-intervention RCADS Total score represented a z-score above 3. Furthermore, analysis of the box plots (see section 3.7.1 for details on data screening) showed that the parent reported pre-intervention RCADS Total score was an outlier, while the parent reported post-intervention RCADS Total score was an extreme outlier for this child.

When the data set for this particular child was removed and the analysis of the parent informed total RCADS data re-run (see Appendix 36), there was no main effect of timepoint, treatment or gender. Neither IQ nor Age were significant covariates. There were no significant 2-way or 3-way interactions.

9.3.1.2 RCADS Total Score Self-Report

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the RCADS Total score when informed by the child revealed that there was no main effect of timepoint, treatment or

gender (see Appendix 36). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions.

9.4 SNAP-IV – The Effect of Dietary Intervention

SNAP-IV data were not completed by the children and were analysed separately by informant (teacher and parent). Table 19 summarises these findings

9.4.1 SNAP-IV Inattention Sub-scale

9.4.1.1 SNAP-IV Inattention Sub-scale Teacher Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SNAP-IV Inattention sub-scale score when informed by teacher revealed that there was no main effect of timepoint or treatment (see Appendix 36). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that girls were reported lower SNAP-IV Inattention sub-scale scores by their teachers compared to boys (mean difference-0.91 ± 0.22, $p < 0.001$). IQ was not a significant covariate. Age was a significant covariate such that age was positively related to the SNAP-IV Inattention sub-scale score with the score increasing as age increased (Figure 21). There were no significant 2-way or 3-way interactions.

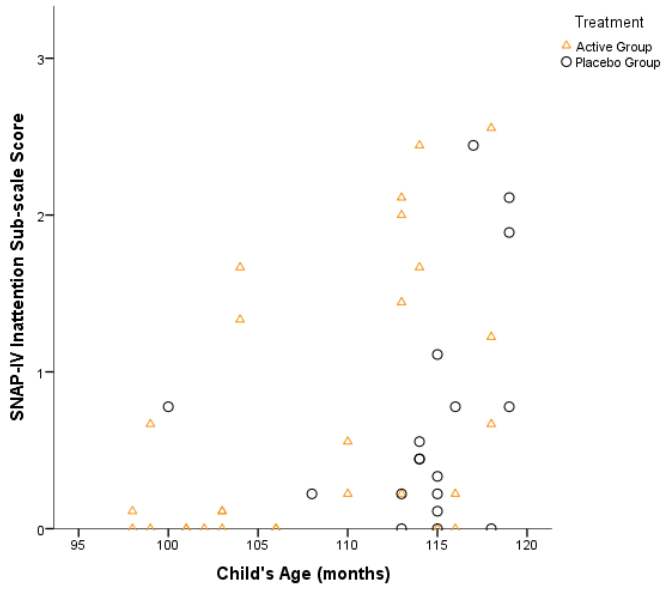


Figure 21: Scatterplot depicting child's age against SNAP-IV Inattention score when informed by the teacher, labelled by treatment group

9.4.1.2 SNAP-IV Inattention Sub-scale Parent Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SNAP-IV Inattention sub-scale when informed by parents revealed that there was no main effect of timepoint or treatment (see Appendix 13). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that the parents of girls their children as having lower Inattention scores on the SNAP-IV compared to parents of boys (mean difference -0.78 ± 0.23, $p \leq 0.001$). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions

9.4.2 SNAP-IV Hyperactivity Sub-scale

9.4.2.1 SNAP-IV Hyperactivity Sub-scale Teacher Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SNAP-IV Hyperactivity sub-scale score when informed by teachers revealed that there was no main effect of timepoint or treatment (see Appendix 36). However, there was a significant main effect of gender. Bonferroni corrected post hoc analysis showed that the parents of children girls rated their children as having lower Hyperactivity scores on the SNAP-IV compared to parents of children boys (mean difference -0.87 ± 0.21, $p < 0.001$).

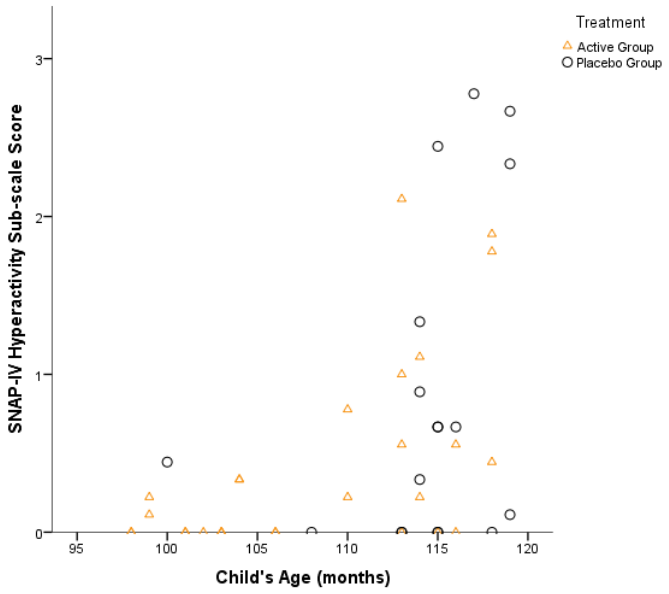


Figure 22: Scatterplot depicting child's age against SNAP-IV Hyperactivity score when informed by the teacher, labelled by treatment group

IQ was not a significant covariate. Age was a significant covariate such that age was positively related to the SNAP-IV Hyperactivity sub-scale score with the score increasing as

age increased (Figure 22). There were no significant 2-way or 3-way interactions, apart from the interaction between treatment x gender (Figure 23). These data are pooled across both pre and post time points so it could be that boys with increased levels of hyperactivity were allocated to the placebo group by chance (although this is perhaps unlikely given random allocation). Boys receiving the Placebo juice were rated as having increased Hyperactivity compared to those in the Active group, but there is no difference between treatment groups for the girls.

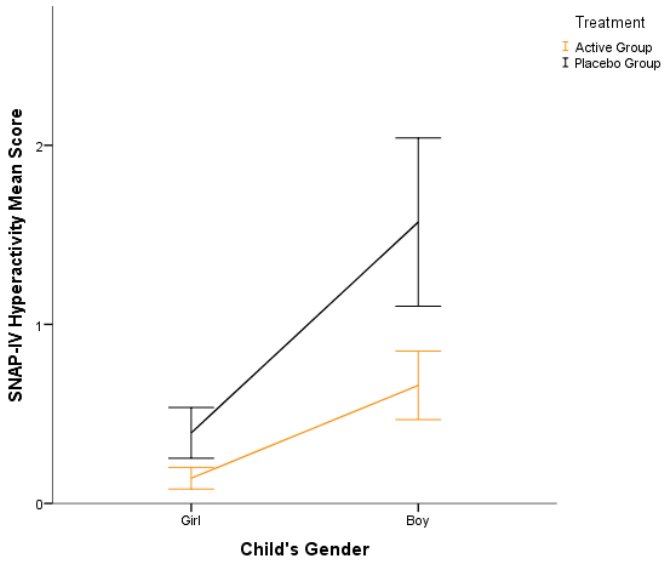


Figure 23: Mean (+/- S.E) depicting child's gender against SNAP-IV Hyperactivity score as informed by the teacher, labelled by treatment group

9.4.2.2 SNAP-IV Hyperactivity Sub-scale Parent Responses

Following the backwards-stepwise procedure for model selection, ensuring the smallest possible AICc, no interactions were removed. Analysis of the SNAP-IV Hyperactivity score when informed by parents revealed that there was no main effect of timepoint, treatment or gender (see Appendix 36). Neither IQ nor age were significant covariates. There were no significant 2-way or 3-way interactions.

Chapter Ten: Discussion – Exploring the Effect of Dietary Intervention on Behaviour and Wellbeing

10.1 Overview of the Thesis Aims

In the development of the tools used within this thesis, it was demonstrated that wellbeing and behaviour difficulties can be measured in young schoolchildren aged 8 to 10 years, using the self-report version of the SDQ. However, the data generated by self-report in this young sample did not fit the proposed 5-factor structure. The emerging consensus (Di Riso et al., 2010; Koskelainen et al., 2001; Percy et al., 2008; Ruchkin et al., 2008) in support of a three factor model for the SDQ, as found in the current thesis, suggests that children understand the items (and the constructs of wellbeing and behaviour) somewhat differently from their parents and teachers. Therefore, it may be advantageous and appropriate to use the SDQ for children in combination with versions completed by other informants (i.e., parents and teachers).

Given the suggested links between polyphenols, behaviour and mood, this thesis aimed to explore the effects of flavonoid consumption on reports of behaviour and wellbeing in young children. The study employed a randomised chronic, 10-week, parallel group, double blind, placebo controlled design to investigate the effects of daily consumption of a flavanone-rich orange juice drink, compared with an equicaloric placebo drink (orange flavoured drink), matched for vitamin C content, on behaviour and wellbeing of a sample of 8 to 10 year old schoolchildren, as reported by multiple informants (child, parent and teacher) using a number of measures, which were explored for their psychometric properties and appropriateness. These measures were (1) the RCADS, a measure of anxiety and depression, (2) the SNAP-IV, a measure of ADHD behaviours and (3) the SDQ, a general screening tool assessing difficulties and strengths.

10.2 Summary of Thesis Results

10.2.1 Effect of Treatment

There was a significant effect of treatment on both the SDQ sub-scales (Total Difficulty and Prosocial Strengths) when reported by the parents. Parents whose children were in the Active group reported their children to have lower Total Difficulty scores and higher Prosocial Strengths, compared to parents whose children were in the Placebo group, when data were averaged over both timepoints (pre and post).

No treatment x time interactions were found (1) reports of behaviour and wellbeing difficulties, as represented by the SDQ Total Difficulty sub-scale and (2) reports of strengths, as represented by the SDQ Prosocial Strengths sub-scale when informed by the parents, teachers or the child themselves, alongside (3) reports of anxiety and depression, as represented by the RCADS Total score when informed by parents or the child and (4) reports of behaviours associated with ADHD, as represented by the SNAP-IV sub-scales (Inattention and Hyperactivity) when informed by parents and teachers. This is inconsistent with suggestions that specific dietary inventions can affect measures of mood. For example, a series of publications has suggested that a formula of 36 vitamins and minerals may affect mood, wellbeing and behaviour of children (Kaplan, Crawford, Field, & Simpson, 2007). Two case studies included in Kaplan's review found a benefit of 2-year micronutrient supplement on behaviour and wellbeing, assessed using items from the Conners Parent Rating Scale, as well as the Child Behavior Checklist (Kaplan, Crawford, Gardner, & Farrelly, 2002). Similarly, Kaplan, Fisher, Crawford, Field, and Kolb (2004) found benefits of the same intervention over 8-weeks on behaviour and wellbeing, assessed with parent-informed versions of the Child Behavior Checklist (CBCL), Youth Outcome Questionnaire (YOQ), and Young Mania Rating Scale (YMRS).

There was an interaction between treatment and IQ on the SDQ Prosocial Strengths sub-scale when informed by the teachers, parents and the child themselves. Children within the Placebo group with higher IQ scores were rated by their teacher as having lower Prosocial Strengths, compared to those in the Active group. The reverse was true when reported by parents, such that parents whose children were in the Placebo group with lower IQ scores were rated as having lower Prosocial Strength scores, while children in the Active group with lower IQ scores were rated by their parents as having more Prosocial Strengths. This latter finding is consistent with findings from the original Goodman (1997) paper which reports a significant positive correlation between parent reported Prosocial Strengths and IQ (specifically for those children with IQ < 70), such that as IQ score increased as did parents reported higher Prosocial Strength scores. This could reflect the link between IQ and conduct problems (thought to be the opposite construct to Prosocial Strengths), whereby low IQ can be a consequence of behavioural difficulties during the development of reading (Adams, Snowling, Hennessy, & Kind, 1999), an effect potentially ameliorated in the Active group. However, no treatment by time interaction can confirm this. The treatment by IQ interaction was also evident for reports of behaviour and wellbeing as represented by SDQ self-report, using the three-factor structure producing sub-scales Extroversion and Hyperactivity. However, it is important to note that the IQ explained only 19% of the variance in the placebo group, with no different seen for children in the Active group.

There was an interaction between treatment and age on the SDQ Total Difficulty composite score when informed by the parents. Younger children receiving the Placebo juice were rated as having more difficulties compared to those in the Active group. However, there was no difference in Total Difficulty score between older or younger children the Active group. This pattern is similar to that seen with the Total RCADS score; however, with the data from the SDQ Total Difficulty composite score, the removal of the outlier did not remove the significant 2-way interaction. Nevertheless, it is likely that this effect is driven by three data points influencing the regression lines (seen in Figure 14), especially in the context of the small sample size used in this thesis. Conversely, it is possible that the polyphenols are having an ameliorating effect at a younger age. However, no treatment by time interaction can confirm this.

10.2.2 Main Effects of Time, Age and Gender

10.2.2.1 Effect of Time

The results showed no effect of time (pre- / post- intervention) on the SDQ or any of the measures employed. The analysis also failed to show an effect of time (pre- / post-intervention) on any of the three newly derived sub-scales of the SDQ (extroversion, introversion or hyperactivity; see section 6.2.4.3). There were no significant treatment by time interactions on any sub-scales across any of the measures completed by any of the informants.

Although preliminary, these results suggest that there is no relationship between flavonoids and wellbeing and behaviour reflected by lack of significant interaction effects of treatment by time on the SDQ Total Difficulty and Prosocial Strengths sub-scale, the RCADS Total sub-scale and the SNAP-IV sub-scales.

This is consistent with the few studies that have examined the effects of a polyphenol-rich orange juice intervention on behaviour and wellbeing. Kean et al. (2015) found no significant treatment effects on measures of anxiety or depression using the HADS, following an 8 week intervention in a sample of healthy older adults. Similarly, Alharbi et al. (2016) found no significant effects of an acute (6 hour) intervention on mood ratings, in a sample of healthy middle-aged adult males.

This is not consistent with studies which have found that the *acute* consumption of a flavonoid-rich drinks can increase positive affect, two hours following consumption, in healthy young children (aged 7 to 10 years old: Khalid et al., 2017). The current findings underline the lack of support for a *chronic* effect of a flavonoid-rich dietary intervention, as

the findings are consistent with a chronic study by Kean (2015), who found subjective mood ratings in healthy older adults to be unrelated to 8-week daily flavonoid consumption.

The non significant effect of treatment by time found in the current study, in contrast to the significant effects found in acute studies, could be explained by the sugar content of the drinks, as opposed to flavonoid concentration. While the sugar content of the Active group and the Placebo group were matched within the current study, sugar content has been shown to be associated with acute (but not chronic) effects on mood (Benton & Nabb, 2003; Sayegh et al., 1995). In contrast, although several correlational studies have suggested that sugar also plays a role in hyperactivity, the experimental data does not support this association (Bellisle et al., 1998).

Alternantivley, Benton & Roberts (1988) suggested that specific nutritional deficiencies (particularly of iron, zinc, and vitamins A and C) occur commonly in British schoolchildren. The study found that when the children (aged 12 to 13 years) had taken vitamin supplement for eight months, those taking the active supplement had significantly higher scores on non-verbal intelligence. While the vitamin C content of the Active group and the Placebo group were matched within the current study, suggesting that this may have masked treatment effects. Yet, significant time effects might have been expected with any effect of vitamin restoration on measures of wellbeing and behaviour.

Although the evidence for the effect of omega-3 on behaviour and wellbeing is mixed in both community (Giles, Mahoney, & Kanarek, 2013) and ADHD samples (Bloch & Qawasmi, 2011; Gillies, Sinn, Lad, Leach, & Ross, 2012). A study using the Child Behaviour Checklist (CBCL) found significant long-term effect for both externalizing and internalizing problems (as reported by parents), after 6 months omega-3 supplementation in a community-residing sample of children (8 to 16 year olds).This may suggest that longer treatment period may be more effective in producing longer-term brain and behavioral change (Raine, Portnoy, Liu, Mahoomed, & Hibbeln, 2015)

The SDQ has been used extensively in intervention research, and specifically used to assess the relationship between diet and wellbeing in children and adolescents (O'Neil et al., 2014). A previous study by Kirby, Woodward, Jackson, Wang, and Crawford (2010) found a significant effect of time, following a 16-week omega-3 intervention with 450 healthy children aged 8 to 10 years old, on the SNAP-IV Inattention and Hyperactivity sub-scales, such that teachers reported lower scores at follow-up compared to baseline (independent of treatment group). They also found that teachers reported lower scores on SDQ sub-scales (Total Difficulty, Emotional Difficulties, Conduct Difficulties, Hyperactivity) after 16 weeks compared to baseline (when controlling for baseline scores), demonstrating the

sensitivity of these measures to change over time and in response to treatment. In the absence of an effect of treatment, an effect of time (pre- and post-) permits the examination of the test-retest reliability of the measures in use.

10.2.2.2 Main effect of Age

There was a significant effect of age on both the SDQ sub-scales (Total Difficulty and Prosocial Strengths) and both SNAP-IV sub-scales (both Inattention and Hyperactivity) when informed by adults. Parents reported higher SDQ Total Difficulty and higher Prosocial Strengths scores for older children, and teachers reported higher SDQ Total Difficulty scores alongside higher SNAP-IV Inattention and Hyperactivity scores, albeit in a relatively narrow range examined in the current study. Consideration of the mean SDQ Total Difficulties scores from previous research (see summary of studies in Appendix 30) suggests that this findings is inconsistent with the trend observed by research (Du et al., 2008; Koskelainen et al., 2000; Meltzer et al., 2000) for adults to report lower SDQ Total Difficulties for older children. Furthermore (as discussed in section 6.2.1.2), language deficits in younger children are thought to be associated with increased reports of hyperactivity and behavioural problems (in a population of 4 to 6 year olds: Norbury et al., 2016; Woerner et al., 2004). An alternative view might suggest that older children have the ability to communicate their internalised difficulties (assessed by the Emotional Difficulties and Peer Problems sub-scales on the SDQ) to their parents and teachers, which is driving this effect on the Total Difficulty composite score. However, these sub-scales were not included in the analysis plan to reduce the number of tests performed in order to reduce the probability of a type-1 error. Nevertheless, this finding is consistent with the data collected when developing the tools for the current thesis. Overall, the effect of age is that more problematic behaviour is reported for older children on the SDQ when reported by parents and teachers.

10.2.2.3 Main Effects of Gender

Gender affected reports by parents, teachers and children on a number of the SDQ outcomes (also discussed in section 6.2.1.3). All measures reported by teachers showed an effect of gender, such that teachers reported higher difficulties scores for boys compared to girls. This effect was also seen for SDQ Total Difficulty when reported by the parent and the child. It may be that differences in the SDQ Hyperactivity sub-scale are driving this significant effect on the SDQ Total Difficulty composite score. However, this sub-scale was not included in the analysis plan to reduce the number of tests performed in order to reduce the probability of a type-1 error. Nevertheless, boys report themselves to have higher scores on SDQ Extroversion and Hyperactivity sub-scales. Complementary to this, is the significant effect of gender is also seen on the SNAP-IV sub-scales (both Inattention and

Hyperactivity), such that girls were more likely to be rated lower on the SNAP-IV sub-scales than boys, by both types of informant (the teacher and parent). This is consistent with findings from Ullebø, Posserud, Heiervang, Obel, and Gillberg (2012) that teachers reported higher scores of Hyperactivity and Impulsivity (and the combined symptom constellation) on the SNAP-IV for boys compared to girls, in a sample of 7 to 9 year old children living in Norway. However, within the Norwegian study difference between genders was less marked according to parent reports. The current study also found significant gender effects for SDQ Prosocial Strengths sub-scale when reported by the teachers, such that they reported higher Prosocial Strength scores for girls than for boys. It is thought that this may reflect the perception that Prosocial Strengths are thought to be the opposite construct to behavioural difficulties such as Hyperactivity. No significant treatment by gender interactions were found.

Although, gender did not affect response to treatment, a gender-specific effect may be somewhat consistent with the current literature. Gender specific effects to dietary intervention have been showed in other studies, for example, Trebatická et al. (2006) found significant effects of Pycnogenol® (a standardised extract of French maritime pine bark containing mixture of natural polyphenols) in 6 to 14 year old girls but not in boys. Nevertheless, the authors dismissed this gender-specific effect due to the small sample size of the study. There is some evidence of gender differences in the physiological response to food (Benton, Haller, & Fordy, 1995). For example, Cook and Benton (1993) found gender differences in the association between mental health and the consumption of fruit and vegetables. However, there is currently no evidence to suggest a gender-specific mechanism underlying the effects of flavonoids on the brain (Khalid et al., 2017).

10.2.2.4 Overall Summary of the Thesis Results

Overall, the intervention study presented in this thesis failed to detect any effect of a 10-week flavonoid-rich dietary intervention compared with a placebo drink matched for calories and Vitamin C content on measures of behaviour and wellbeing in a small sample of healthy young children based on teacher, parents and the child's reports. This conclusion applies across all of the measures employed.

10.3 Strengths of the Research presented in this Thesis

The findings of this thesis can be considered important for a number of reasons. Namely, to the author's knowledge, the work presented in this thesis is novel in its development of appropriate tools through the exploration of the SDQ factor structure based on different informants to assess the behaviour and wellbeing of 8 to 10 year old UK schoolchildren.

This is also the only study investigating the effect of a chronic (10-week) flavonoid-rich orange juice dietary intervention on behaviour and wellbeing, using multiple informants, in a UK sample of healthy children aged 8 to 10 years.

10.3.1 The Use of Cross Informants

A major advantage of the research presented in this thesis is the use of cross-informant comparisons; obtaining scores from parents, teachers and the children themselves. Informants varied in the extent to which they reported different types of behaviours and predictions (in terms of diagnosis or outcome) are more likely to be accurate when based on triangulation of multiple informants (Kraemer et al., 2003). For example, behavioural disorders may be more apparent in certain settings (e.g school rather than at home) and the teacher may be a more objective observer of such behaviours based on their training and experience of a range of children. Children in turn, can offer additional information about their worries, anxieties and behaviour that perhaps parents and teachers were not aware of (Goodman, 2000).

10.3.2 Implications of the research presented in this Thesis

The findings have important implications for the use of self-report measures, with young people, in clinical practice. The factors extracted for the children's self-report of their behaviour and wellbeing tell us that they are conflating Prosocial Strengths, Peer Problems and Emotional and Conduct Difficulties into two binary categories, seemingly driven by social desirability. The new three-factor structure elucidated in this thesis could improve the reliability of the self-report version when used as a general screening instrument (serving as a baseline of a children's emotional and behavioural state) for young children (aged 8 to 10 years old) in community settings (e.g. schools). When used alongside the adult informant versions as a routine outcome measure, the SDQ self-report version may promote early identification of mental health difficulties, leading to early intervention and may facilitate development of the evidence-base for interventions with young children. Given the current economic climate, a screening tool that is brief and inexpensive to administer and analyse may be considered advantageous to service providers. The current findings support claims that when only one informant is available, reports from teachers are a better choice than the use of SDQ self-report in children alone (Aitken, Martinussen, & Tannock, 2017; Goodman, Ford, Corbin, & Meltzer, 2004; Goodman et al., 2003).

The findings also have important implications for the exploration of the link between polyphenols and behaviour and wellbeing. The current findings suggest that further work is

required to understand the active dosage and duration of polyphenols required to detect effects on wellbeing and behaviour in both community and clinical samples of children.

10.4 Limitations of the Research presented in this Thesis

However, there are various factors which influence the outcomes of studies even when they are well controlled for (as within the current study), therefore, the following issues need to be considered when interpreting the findings.

10.4.1 Consideration of the Polyphenol Treatment used

It is important to point out that the dose of polyphenols used within the current study is less than that used in comparable studies that found effects of polyphenol-rich dietary interventions on mood. For example, Khalid et al. (2017) found that *acute* consumption of a flavonoid-rich blueberry drink containing 253mg anthocyanins improved positive affect after 2 hours in a community sample of 7 to 10 year old children. While, Alharbi et al. (2016) found that acute consumption of flavonoid-rich orange drink containing 272mg of flavonoids did not significant effect mood in a community sample of *adults*. In a *chronic* study, Trebatická et al. (2006) found that Pycnogenol® at a daily dose of 1mg/kg of child's body weight (estimated between 26–32kg for 8 to 10 year old children, according to UK Child Growth Foundation Data, cited in: Carasco, Fletcher, & Maconochie, 2016) over a one month period improved measures of ADHD behaviours in children with ADHD. In contrast to these examples, the dosage of polyphenols sufficient to detect cognitive benefits can be lower. Research suggests that a daily dose of polyphenol-rich supplement containing a total of 740mg of polyphenolics over a 12 week period is appropriate to detect polyphenol-induced changes on measures of cognition (Lamport, Pal, et al., 2016). The current study used approximately 120mg per 8oz bottle per day, suggesting that this dose may not have been sufficient to affect measures of mood and behaviour. It may also be relevant to note that, there is a small possibility that the post-intervention data collection was confounded by acute effects (see Figure 12: Study Schedule Schematic). However, as discussed above, studies that have found acute effects of polyphenols on measures of mood and behaviour (Khalid, et al., 2017) have used much higher doses than those used in the current study. Therefore any effects were considered to be a result of a cumulative effect of polyphenols (Lamport, Lawton, et al., 2016) across the 10-week intervention. Interestingly, only one dose was used in the current study, whereas studies looking at the effect of polyphenols on cognitive performance have found significant effects of dose (Scholey et al., 2010) and have recommended the use of multiple doses and time points to shed more light on the impact of flavonoid intervention on performance (Whyte & Williams, 2015).

Correspondingly, although compliance to the treatment was monitored, we cannot know that the child definitely consumed the full dose. Compliance was observed in the classroom and a drink diary log sheet was included for weekends (on which parents or children could tick off each day's drink) to monitor compliance. On-the-other-hand, ensuring consumption on school days should have ensured a reasonable dose, as this ought not to have washed out over the weekend/holidays/sick period. While, early studies suggest that here are not expected to be any carry-over effects of flavonoids, as the half-life of flavonoids is estimated to range from 2–28 h (Manach & Donovan, 2004), in contrast, a cumulative effect of polyphenols has been suggested, such that there is a carryover effect of polyphenols even with long washout periods (Lamport, Lawton, et al., 2016). As the current study is exploring a chronic cumulative effect of polyphenols on behaviour and wellbeing, compliance may not be considered a limitation.

Given the proposed cumulative effects of polyphenols (Lamport, Lawton, et al., 2016), it is relevant to note that this study did not include extensive dietary measures, nor did it control for habitual consumption of foods which may confound the results. For example, there was no assessment of the intake of fruit and vegetables in the children's normal diet, which may have influenced the effects of the intervention. This is particularly relevant as the literature suggests that successfully dietary interventions are often associated with nutritional deficiency (Falkingham et al., 2010), perhaps, for example, as a result of low fruit and vegetable intake. However, the Index of Multiple Deprivation (IMD) data for the recruitment area places the sample across a range of deprived deciles (2-9th deciles: CDRC, Accessed 2018). Therefore, in contrast, this could also have be influenced by the nature of a self-selected sample, it is possible that parents willing to engage in research are also likely to feed their children more fruit and vegetables. A diet high in fruit and vegetables is, therefore, high in polyphenols which may mean that a polyphenol-rich dietary intervention is not adding something new or correcting a nutritional deficiency (this problem is commonly referred to as a ceiling effect).

10.4.2 Impact of the Sample-size

One of the more obvious limitations was the sample size and the high drop-out rates of adult informants (parents and teachers). It is likely that the results were biased by the loss of data that was expected to be completed by the adult informants (parents and teachers), especially at follow-up of the intervention study (for discussion of selection Bias see section 6.3, and for discussion of non-random missing data from teacher and parents see section 6.4). Given the current constraints on the education system difficulties were encountered when attempting to collect data from the teachers, which will have reduced the statistical power

and increased the probability of type II error, which may explain the lack of an effect of the intervention

It is an extremely difficult and time-consuming task to set up this type of school-based group intervention. It is therefore essential that a dedicated teacher or member of school staff is involved in the planning and execution of the study. It is difficult to meet with the children to explain the purpose of the study, collect signed parental consent forms, perform screening assessments prior to the first session and, importantly, distribute the daily intervention without the help of a very motivated member of school staff, whom the children know and trust. Therefore, this study was only able to recruit small numbers of children from two local schools and there were only a small number of participants who fully completed the study protocol. The key reasons (anecdotally provided by the children) for not participating was the high demand of the study as whole (including the microbiota data collection; data not included in the present thesis) or that the dietary supplement (the orange juice) was not liked. The impact of this on the sample size will likely have reduced the statistical power and increased the probability of type II error. The replication of this research on a larger scale would clarify whether the lack of effect observed is a true lack of effect or a type II error.

10.4.2 Consideration of the Sensitivity of the Sample and the Measures used

The research sample consisted of healthy children in mainstream school with minimal to mild baseline scores on all the measures of behaviour and wellbeing (see section 9.1). It is perhaps, therefore, not surprising that a beneficial effect of the flavonoid-rich intervention was not detected on these outcomes. This observation is consistent with the findings reported by Benton, Williams, and Brown (2007), who found that improvements in mood after dietary interventions (for example, probiotics) only occurred in participants who showed elevated symptoms of depression at the baseline (this problem is commonly referred to as a floor effect). It is not uncommon in the literature that findings do not show intervention effects on parents and teacher ratings of children's behaviour because base rates of these behaviours as low (Muratori et al., 2017; for discussion around informant biases see section 5.2.1.1). This could also have been influenced by the nature of a self-selected sample, parents may be less likely to respond if their child is experiencing difficulties. This might be reflected in the relatively larger levels of non-completion for the SNAP-IV, an ADHD screening tool, which asks direct questions about undesirable problematic behaviour. Although the use of a nonclinical sample provided the opportunity to assess the effects of the intervention on these measures independent of the confounding

effects of ongoing behavioural and emotional difficulties, further research could account for these concerns.

Even though the work reported in this thesis did not find an effect of flavonoid-rich orange juice on any measure of behaviour and wellbeing, there is some suggestion that flavonoids are specifically associated with depression, in contrast to anxiety or other mood states (see section 6.4). Thus, the findings of this thesis may reflect the use of a non-specific measure of behaviour and wellbeing. It is relevant to note that the current study found no effect of the polyphenol-rich intervention on the RCADS Total score (reported by either parent or the child themselves), but this composite score does not distinguish between anxiety and depression symptomology. Due to the nature of measuring behaviour and wellbeing, the lack of a significant effect of flavonoid-rich orange juice on the measures used in this study does not rule out the possibility of benefits on behaviour and wellbeing using more specific measures (such as the RCADS Depression sub-scale). The current study chose to analyse only the primary dependent variables of interest (SDQ Total Difficulty composite score and Prosocial Strengths sub-scale score, RCADS Total Score and SNAP-IV Inattention and Hyperactivity sub-scale scores). Previous studies have used batteries' of multiple measures to detect an effect of polyphenol-rich supplements on cognition and mood, perhaps increasing the likelihood of encountering a type 1 error. It is also important to acknowledge that non-significant reductions in measures of behaviour and wellbeing may not necessarily imply that important changes did not occur. Even small improvements in behaviour and wellbeing may be personally relevant to individuals and their families.

All the participating children were living in a semi-rural, suburban areas which led to an ethnically homogeneous (predominately identifying as "white") sample. The little variation in ethnicity could have introduced bias and the generalisability of the current findings to other groups should be questioned. This may be relevant as there is some evidence to suggest that children living in large cities have more behavioural problems and these children were recruited from schools in semi-rural outlying villages (Larsson & Frisk, 1999). Socioeconomic status is a complex but potentially important confounder given it is association with both diet and mental health in adults (Brennan, Henry, Nicholson, Kotowicz, & Pasco, 2010; Williams et al., 2011). The high number of factors that could confound a wellbeing and behaviour measure, in contrast to a more objective measure of cognition, may explain some of the differences between the current findings and the polyphenol literature base. Further studies might, therefore, be advised to measure social, educational and economic factors that may influence children's mood. For example, Ford, Macdiarmid, Russell, Racey, and Goodman (2017) found conduct disorder and ADHD were predicted by factors such as rented housing, large family size, and poor family function as well as by poor peer relationships and intellectual ability scores. As another example, the

current research might have considered the mediating role of parental behaviour on the child's behaviour and wellbeing. A study by Raine et al. (2015) found parental improvement in antisocial behaviour partly mediated improvements in child behaviour produced by omega-3 supplementation. It is, perhaps, not surprising to anticipate improvements in child behaviour evoking improvements in parental behaviour that further improves child behaviour, given that caregivers and their children are thought to have reciprocal effects on one another's behaviours (Dodge & Pettit, 2003).

A number of the limitations are shared with all epidemiological studies employing a questionnaire-based methodology. Importantly, no attempts were made to adapt the SDQ for use with a younger sample. With regard to content, Marsh (1986) found that children (aged 7 to 12 years old), specifically younger children and those with poor verbal skills, were less able to respond to negatively worded items. Consider also, the use of the three-point Likert-style rating scale (used in the current research): on-one-hand, the use of a rating scale approach can lack specificity or depth, as the Likert response technique allows informants to choose from only three options, which are coded from zero to two. On-the-other-hand, limited choices have found to improve reliability in younger samples. Mellor and Moore (2013) promote the dichotomous yes/no rating as the gold standard, as they believe that it provides the least ambiguity. However, this approach would be less sensitive to change and a larger switch from negative to positive would be required to demonstrate significant change. The recent study by Mellor and Moore (2013) emphasized the importance of anchoring the Likert-scale with words that are meaningful to children. It was suggested that the easiest format for children appears to be that which is based on words that reflect frequency of behaviors/thoughts (i.e., never to regularly). In consideration of the capacity of children to respond to such scales, there could have been more awareness of this by: carefully choosing age appropriate reading level vocabulary, reducing the number of response choices, or using visual aids (such as faces on which various degrees of happiness or sadness are depicted).

10.4.3 Consideration of the Ecological Validity

A further factor is that testing took place in a school setting, which is associated with an unavoidable level of noise distraction. This represented a more ecologically valid approach, whilst an alternative might have been to conduct the study in a more scientifically controlled environment, to ensure greater compliance with study procedures. This setting may also allow for additional measures (such as blood measures of flavonoid and metabolite content in an adult sample, but would not be feasible in children).

10.5 Future Research Directions informed by this Thesis

A number of future research directions are suggested by the results presented within this thesis. Firstly, sample size issues and methodological concerns raised could be accommodated in a replication study in a more diverse and larger sample of children. This may consist of the inclusion of an older control group of children (aged 11 to 16 years old) and an adapted version of the SDQ (to make it more appropriate to the reading age of 8 to 10 year old children). It may also be beneficial to perform qualitative analysis of the individual items to determine if there are any differences in understanding between older and younger children, and to further clarify the mechanisms behind the discrepancies that are often reportedly found between self- and informant-reports.

More specifically, further studies might also be advised to measure background diet (Lamport, Lawton, et al., 2016), especially dietary intake of polyphenol rich foods and take this into consideration in the data analysis (e.g. by splitting participants into "high" or "low" fruit and vegetable consumption groups) or when screening participant (e.g. set a limit on fruit and vegetable intake as an inclusion criterion, bearing in mind that it would be unethical to ask participants to refrain from consuming fruit and vegetables in the study).

Although the work reported in this thesis does not focus on biological mechanisms that could underlie possible beneficial effects of flavonoid-rich dietary interventions on behaviour and wellbeing, the literature does offer a number of hypotheses which may be testable in future studies. One of these hypotheses relates to the role of intestinal microbiota[1] (or gut microbiota) as a possible mediator between diet and psychological functioning which has been the subject of recent speculation (Cani et al., 2007; Cryan & Dinan, 2012; Dinan, Stanton, & Cryan, 2013; Knowles & Cole, 2008).

Finally, it may be of benefit to further explore the link between polyphenols and behaviour and wellbeing, using a more specific measure while more carefully considering factors such as the dose and confounding variables in a context specific to wellbeing and behaviour.

10.6 Summary of Discussion

Overall the work presented in this thesis does not support a beneficial effect of 10-weeks consumption of flavonoid-rich orange juice drink on the behaviour and wellbeing of 8 to 10 year old UK schoolchildren. However, it was suggested that the outcome measures used may not have been sufficiently sensitive to detect the small changes produced by this small

[1] Gut microbiota was assessed within the main RCT, however it is beyond the scope of this thesis to consider the role of gut microbiota.

dietary intervention. In addition, missing informant data, particularly for parents and teachers compromised the analysis by reducing power. Therefore, further work is required to substantiate the current findings in the light of the methodological limitations of the current study. In particular, it would be beneficial to repeat the study with a larger sample size whilst controlling for confounding factors such as habitual intake of flavonoids. Further work would also be strengthened by concurrent assessment of biological endpoints which might shed light on the mechanism of action of any effects of flavonoid and other dietary interventions on behaviour and wellbeing.

Chapter Eleven: Overall Conclusions of the Thesis

Recent national and international policies recommend the use of patient reported outcomes across all health services, including CAMHS. The assessment of behaviour and wellbeing of children has traditionally relied on adult informants, while the focus on client collaboration is promoting the importance of the voice of the child. Although the measurement of these constructs in children poses some difficulties, there are a number of valid and reliable measures currently in use. This thesis focused on three measures: (1) SDQ, a general measure of behaviour and wellbeing, (2) RCADS, a measure of anxiety and depression, and (3) SNAP-IV, a measure of ADHD symptomology. However, the vast majority of research on self-reported measures focuses on older children and adolescents. Yet in practice, these measures are often used with younger populations. It is important that these rating scales provide meaningful information for service users, clinicians and the services. Clinicians should ensure that they are using these tools appropriately, taking into account the psychometric properties of the measures distributed to specific populations, to determine the relevance to the clients and ability of the client to complete it accurately. Therefore, in developing appropriate tools for this thesis, the value of self-report measures with young children (aged 8 to 10 years old) was established, in the accurate reflection of a child's behavioural and emotional difficulties. This was achieved by exploring the psychometric properties, reliability and validity of SDQ data collected from different informants; parents, teachers and the children themselves.

Overall, the findings of this thesis suggest that the teacher informant responses are perhaps with most useful in assessing the behaviour and wellbeing of young children but that these reports can be challenging to obtain for research purposes. Firstly, as a homogeneous group, the teachers demonstrate greater internal consistency (discussed in section 5.2.2.2), suggesting that measures are more likely to be addressing the same construct between raters. Furthermore, the proposed five-factor structure was evident within the SDQ data when completed by the teachers. Regarding the self-report version, the results using the five-factor structure showed adequate internal consistency, but poor agreement with the adult informants and poor concurrent validity. The exploratory factor analysis confirmed that the five-factor structure was not the best fit for this sample of informants. A new three-factor structure improved the reliability of the measure. The study suggested that the SDQ self-report version, although designed for children aged 11 years and over, can provide valid information about behaviour and wellbeing when informed by children aged 8 to 10 years old, provided that it is scored appropriately using a factor structure that better signifies the responses of younger children. This thesis has furthered the understanding of self-report

measures, when used in younger children, and emphasized the value of psychometric consideration of these tools when working clinically.

Following the development of appropriate tools, the thesis explored the links between diet, cognition, behaviour and wellbeing, comparing the reports from different informants; teacher, parent and self-reports of behaviour and wellbeing, using the SDQ, RCADS and SNAP-IV, before and after a 10 week flavanone-rich dietary intervention in a community sample of children aged 8 to 10 years old. It was predicted that the three-factor structure of the SDQ would offer increased validity to the assessment of self-reported changes in behaviour and wellbeing. Although the complex associations between specific dietary interventions and cognitive, behavioural and emotional functioning are not yet fully understood, these preliminary results suggest limited effects of flavonoids on behaviour and wellbeing in a sample of young schoolchildren. However, due to the nature of measuring behaviour and wellbeing, the lack of a significant effect of the orange juice intervention on the outcomes assessed in this research does not rule out the possibility of benefits on behaviour and wellbeing using more specific measures, alternative dosages and while controlling for the complex confounding factors. It is recommended that, due to limitations associated with the sample size, further research should be conducted with a larger sample size to further investigate the effects of polyphenol-rich dietary intervention on behaviour and wellbeing.

The research presented in this thesis highlights the importance of assessing the validity and reliability of measures used to assess the behaviour and wellbeing of young children in clinical practice and research. It also provides an indication of how these measures might be further developed for future work in the field.

References

Achenbach. (1991). CBCL: Child behavior checklist for ages 4-18 [Questionnaire et manuel]. *Burlington: ASEBA*.

Achenbach, Mcconaughy, & Howell. (1987). Child Adolescent Behavioral and Emotional-Problems - Implications of Cross-Informant Correlations for Situational Specificity. *Psychological Bulletin, 101*(2), 213-232. doi: Doi 10.1037/0033-2909.101.2.213

Adams, J. B., Johansen, L. J., Powell, L. D., Quig, D., & Rubin, R. A. (2011). Gastrointestinal flora and gastrointestinal status in children with autism–comparisons to typical children and correlation with autism severity. *BMC gastroenterology, 11*(1), 22.

Adams, J. W., Snowling, M. J., Hennessy, S. M., & Kind, P. (1999). Problems of behaviour, reading and arithmetic: Assessments of comorbidity using the Strengths and Difficulties Questionnaire. *British Journal of Educational Psychology, 69*(4), 571-585.

Adolphus, K., Lawton, C. L., & Dye, L. (2013). The effects of breakfast on behavior and academic performance in children and adolescents. *Frontiers in human neuroscience, 7*, 425.

Adolphus, K., Lawton, C. L., & Dye, L. (2015). The relationship between habitual breakfast consumption frequency and academic performance in British adolescents. *Frontiers in public health, 3*, 68.

Agrawal, R., & Lord, D. (2006). Effects of sample size on goodness-of-fit statistic and confidence intervals of crash prediction models subjected to low sample mean values. *Transportation Research Record: Journal of the Transportation Research Board*(1950), 35-43.

Aitken, M., Martinussen, R., & Tannock, R. (2017). Incremental Validity of Teacher and Parent Symptom and Impairment Ratings when Screening for Mental Health Difficulties. *J Abnorm Child Psychol, 45*(4), 827-837. doi: 10.1007/s10802-016-0188-y

Akbaraly, T. N., Sabia, S., Shipley, M. J., Batty, G. D., & Kivimaki, M. (2013). Adherence to healthy dietary guidelines and future depressive symptoms: evidence for sex differentials in the Whitehall II study–. *Am J Clin Nutr, 97*(2), 419-427.

Alharbi, M. H., Lamport, D. J., Dodd, G. F., Saunders, C., Harkness, L., Butler, L. T., & Spencer, J. P. (2016). Flavonoid-rich orange juice is associated with acute improvements in cognitive function in healthy middle-aged males. *European journal of nutrition, 55*(6), 2021-2029.

Allam, F., Dao, A. T., Chugh, G., Bohat, R., Jafri, F., Patki, G., . . . Salim, S. (2013). Grape Powder Supplementation Prevents Oxidative Stress–Induced Anxiety-Like Behavior, Memory Impairment, and High Blood Pressure in Rats–3. *J Nutr, 143*(6), 835-842.

Alreck, D., & Settle, G. (1985). The survey research book. *Illinois: IRWIN*.

Alwin, D. F., & Krosnick, J. A. (1991). The reliability of survey attitude measurement: The influence of question and respondent attributes. *Sociological Methods & Research, 20*(1), 139-181.

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child neuropsychology, 8*(2), 71-82.

Angelucci, F., Brene, S., & Mathe, A. (2005). BDNF in schizophrenia, depression and corresponding animal models. *Molecular psychiatry, 10*(4), 345.

Assessment, I. E. I. (2010). Liberating the NHS: Greater choice and control.

Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders*.

Baumgartner, J., Smuts, C. M., Malan, L., Kvalsvig, J., van Stuijvenberg, M. E., Hurrell, R. F., & Zimmermann, M. B. (2012). Effects of iron and n-3 fatty acid supplementation, alone and in combination, on cognition in school children: a randomized, double-blind, placebo-controlled intervention in South Africa–. *Am J Clin Nutr, 96*(6), 1327-1338.

Beezhold, B., Radnitz, C., Rinne, A., & DiMatteo, J. (2015). Vegans report less stress and anxiety than omnivores. *Nutritional neuroscience, 18*(7), 289-296.

Beitchman, J. H., & Corradini, A. (1988). Self-report measures for use with children: A review and comment. *J Clin Psychol, 44*(4), 477-490.

Bell, L., Lamport, D. J., Butler, L. T., & Williams, C. M. (2015). A review of the cognitive effects observed in humans following acute supplementation with flavonoids, and their associated mechanisms of action. *Nutrients, 7*(12), 10290-10306.

Bellisle, F. (2004). Effects of diet on behaviour and cognition in children. *British journal of nutrition, 92*(S2), S227-S232.

Bellisle, F., Blundell, J., Dye, L., Fantino, M., Fern, E., Fletcher, R., . . . Westenhöfer, J. (1998). Functional food science and behaviour and psychological functions. *British journal of nutrition, 80*(S1), S173-S193.

Benton, Williams, & Brown. (2007). Impact of consuming a milk drink containing a probiotic on mood and cognition. *Eur J Clin Nutr, 61*(3), 355.

Benton, D., Haller, J., & Fordy, J. (1995). Vitamin supplementation for 1 year improves mood. *Neuropsychobiology, 32*(2), 98-105.

Benton, D., & Nabb, S. (2003). Carbohydrate, memory, and mood. *Nutrition reviews, 61*(s5).

Benton, D., & Roberts, G. (1988). Effect of vitamin and mineral supplementation on intelligence of a sample of schoolchildren. *The Lancet, 331*(8578), 140-143.

Birch, L. L., & Fisher, J. O. (1998). Development of eating behaviors among children and adolescents. *Pediatrics, 101*(Supplement 2), 539-549.

Black, N. (2013). Patient reported outcome measures could help transform healthcare. *BMJ: British Medical Journal, 346*.

Bloch, M. H., & Qawasmi, A. (2011). Omega-3 fatty acid supplementation for the treatment of children with attention-deficit/hyperactivity disorder symptomatology: systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 50*(10), 991-1000.

Bondonno, C. P., Downey, L. A., Croft, K. D., Scholey, A., Stough, C., Yang, X., . . . Swinny, E. (2014). The acute effect of flavonoid-rich apples and nitrate-rich spinach on cognitive performance and mood in healthy men and women. *Food & function, 5*(5), 849-858.

Borg, A. M., Kaukonen, P., Joukamaa, M., & Tamminen, T. (2014). Finnish norms for young children on the Strengths and Difficulties Questionnaire. *Nord J Psychiatry, 68*(7), 433-442. doi: 10.3109/08039488.2013.853833

Bouayed, J. (2010). Polyphenols: a potential new strategy for the prevention and treatment of anxiety and depression. *Current Nutrition & Food Science, 6*(1), 13-18.

Bravo, L. (1998). Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutrition reviews, 56*(11), 317-333.

Brennan, S. L., Henry, M. J., Nicholson, G. C., Kotowicz, M. A., & Pasco, J. A. (2010). Socioeconomic status, obesity and lifestyle in men: the Geelong Osteoporosis Study. *Journal of men's health, 7*(1), 31-41.

Buie, T., Campbell, D. B., Fuchs, G. J., Furuta, G. T., Levy, J., VandeWater, J., . . . Beaudet, A. L. (2010). Evaluation, diagnosis, and treatment of gastrointestinal disorders in individuals with ASDs: a consensus report. *Pediatrics, 125*(Supplement 1), S1-S18.

Bull, R., Lee, K., Koh, I. H., & Poon, K. K. (2016). Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in Singaporean kindergartners. *Child Care Health Dev, 42*(1), 109-116. doi: 10.1111/cch.12288

Bussing, R., Fernandez, M., Harwood, M., Hou, W., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: psychometric properties and normative ratings from a school district sample. *Assessment, 15*(3), 317-328.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81.

Cani, P. D., Neyrinck, A. M., Fava, F., Knauf, C., Burcelin, R. G., Tuohy, K. M., . . . Delzenne, N. M. (2007). Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia, 50*(11), 2374-2383.

Capron, C., Thérond, C., & Duyme, M. (2007). Psychometric properties of the French version of the self-report and teacher Strengths and Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment, 23*(2), 79.

Carasco, C. F., Fletcher, P., & Maconochie, I. (2016). Review of commonly used age-based weight estimates for paediatric drug dosing in relation to the pharmacokinetic properties of resuscitation drugs. *British journal of clinical pharmacology, 81*(5), 849-856.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245-276.

CDRC. (Accessed 2018). Consumer Data Research Centre Maps. Retrieved 15/06/2018, 2018, from https://maps.cdrc.ac.uk

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *J Pediatr Psychol, 27*(1), 27-36.

Chambers, C. T., Reid, G. J., Craig, K. D., McGrath, P. J., & Finley, G. A. (1998). Agreement between child and parent reports of pain. *The Clinical journal of pain, 14*(4), 336-342.

Chang, S.-C., Cassidy, A., Willett, W. C., Rimm, E. B., O'Reilly, E. J., & Okereke, O. I. (2016). Dietary flavonoid intake and risk of incident depression in midlife and older women–3. *Am J Clin Nutr, 104*(3), 704-714.

Chedraui, P., San Miguel, G., & Schwager, G. (2011). The effect of soy-derived isoflavones over hot flushes, menopausal symptoms and mood in climacteric women with increased body mass index. *Gynecological endocrinology, 27*(5), 307-313.

Chorpita, B. F., Moffitt, C. E., & Gray, J. (2005). Psychometric properties of the Revised Child Anxiety and Depression Scale in a clinical sample. *Behaviour research and therapy, 43*(3), 309-322.

Chorpita, B. F., Yim, L., Moffitt, C., Umemoto, L. A., & Francis, S. E. (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: A revised child anxiety and depression scale. *Behaviour research and therapy, 38*(8), 835-855.

Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(9), 1015-1037.

Comrey, A., & Lee, H. (1992). Interpretation and application of factor analytic results. *Comrey AL, Lee HB. A first course in factor analysis, 2*.

Conner, T. S., Brookie, K. L., Richardson, A. C., & Polak, M. A. (2015). On carrots and curiosity: eating fruit and vegetables is associated with greater flourishing in daily life. *British Journal of Health Psychology, 20*(2), 413-427.

Cook, R., & Benton, D. (1993). The relationship between diet and mental health. *Personality and Individual Differences, 14*(3), 397-403.

Cooper, S. B., Bandelow, S., & Nevill, M. E. (2011). Breakfast consumption and cognitive function in adolescent schoolchildren. *Physiology & behavior, 103*(5), 431-439.

CORC. (Accessed 2018). Child Outcomes Research Consortium Retrieved 15/06/2018, 2018, from https://www.corc.uk.net/

Cremeens, J., Eiser, C., & Blades, M. (2006). Characteristics of health-related self-report measures for children aged three to eight years: a review of the literature. *Quality of Life Research, 15*(4), 739-754.

Crichton, G. E., Bryan, J., Hodgson, J. M., & Murphy, K. J. (2013). Mediterranean diet adherence and self-reported psychological functioning in an Australian sample. *Appetite, 70*, 53-59. doi: 10.1016/j.appet.2013.06.088

Crook, W. C. (1974). An alternate method of managing the hyperactive child. *Pediatrics, 54*(5), 656-656.

Cryan, J. F., & Dinan, T. G. (2012). Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature reviews neuroscience, 13*(10), 701.

CYP-IAPT. (Accessed 2018). Children and Young People's Improve Access to Psychological Therapies Retrieved 15/06/2018, 2018, from https://www.england.nhs.uk/mental-health/cyp/iapt/

d'Acremont, M., & Linden, M. V. d. (2008). Confirmatory factor analysis of the Strengths and Difficulties Questionnaire in a community sample of French-speaking adolescents. *European Journal of Psychological Assessment, 24*(1), 1.

D'Souza, S., Waldie, K. E., Peterson, E. R., Underwood, L., & Morton, S. M. (2017). The Strengths and Difficulties Questionnaire: Factor Structure of the Father-Report and Parent Agreement in 2-Year-Old Children. *Assessment*, 1073191117698757.

Davis, C. S. (2002). *Statistical methods for the analysis of repeated measurements*: Springer Science & Business Media.

de Theije, C. G., Wu, J., da Silva, S. L., Kamphuis, P. J., Garssen, J., Korte, S. M., & Kraneveld, A. D. (2011). Pathways underlying the gut-to-brain connection in autism spectrum disorders as future targets for disease management. *European journal of pharmacology, 668*, S70-S80.

Denham, S. A. (1998). *Emotional development in young children*: Guilford Press.

Dex, S., & Hollingworth, K. (2012). Children's and young people's voices on their wellbeing. *Childhood Well-Being Research Centre, Working Paper, 16*.

Di Riso, D., Salcuni, S., Chessa, D., Raudino, A., Lis, A., & Altoè, G. (2010). The Strengths and Difficulties Questionnaire (SDQ). Early evidence of its reliability and validity in a community sample of Italian children. *Personality and Individual Differences, 49*(6), 570-575.

Dickey, W. C., & Blumberg, S. J. (2004). Revisiting the factor structure of the strengths and difficulties questionnaire: United States, 2001. *Journal of the American Academy of Child & Adolescent Psychiatry, 43*(9), 1159-1167.

Diggle, P. (2002). *Analysis of longitudinal data*: Oxford University Press.

Dinan, T. G., Stanton, C., & Cryan, J. F. (2013). Psychobiotics: a novel class of psychotropic. *Biological Psychiatry, 74*(10), 720-726.

Dobbs, J., & Arnold, D. H. (2009). Relationship between preschool teachers' reports of children's behavior and their behavior toward those children. *School Psychology Quarterly, 24*(2), 95.

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology, 39*(2), 349.

Du, Y., Kou, J., & Coghill, D. (2008). The validity, reliability and normative scores of the parent, teacher and self report versions of the Strengths and Difficulties Questionnaire in China. *Child and Adolescent Psychiatry and Mental Health, 2*(1), 8.

Dunckley, M., Aspinal, F., Addington-Hall, J. M., Hughes, R., & Higginson, I. J. (2005). A research study to identify facilitators and barriers to outcome measure implementation. *International journal of palliative nursing, 11*(5), 218-225.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child development*, 265-275.

Einvik, G., Ekeberg, O., Lavik, J. G., Ellingsen, I., Klemsdal, T. O., & Hjerkinn, E. M. (2010). The influence of long-term awareness of hyperlipidemia and of 3 years of dietary counseling on depression, anxiety, and quality of life. *J Psychosom Res, 68*(6), 567-572.

Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., . . . Rollnick, S. (2012). Shared decision making: a model for clinical practice. *Journal of general internal medicine, 27*(10), 1361-1367.

Emerson, E. (2005). Use of the Strengths and Difficulties Questionnaire to assess the mental health needs of children and adolescents with intellectual disabilities. *Journal of Intellectual and Developmental Disability, 30*(1), 14-23.

Epstein, H. T. (1986). Stages in human brain development. *Developmental Brain Research, 30*(1), 114-119.

Falkingham, M., Abdelhamid, A., Curtis, P., Fairweather-Tait, S., Dye, L., & Hooper, L. (2010). The effects of oral iron supplementation on cognition in older children and adults: a systematic review and meta-analysis. *Nutrition journal, 9*(1), 4.

Fallon, T., & Schwab-Stone, M. (1994). Determinants of reliability in psychiatric surveys of children aged 6–12. *Journal of Child Psychology and Psychiatry, 35*(8), 1391-1408.

Feingold, B. F. (1975). Hyperkinesis and learning disabilities linked to artificial food flavors and colors. *The American journal of nursing*.

Ferketich, S. (1991). Focus on psychometrics. Aspects of item analysis. *Res Nurs Health, 14*(2), 165-168.

Ford, T., Macdiarmid, F., Russell, A., Racey, D., & Goodman, R. (2017). The predictors of persistent DSM-IV disorders in 3-year follow-ups of the British Child and Adolescent Mental Health Surveys 1999 and 2004. *Psychol Med, 47*(6), 1126-1137.

Galloway, H., & Newman, E. (2017). Is there a difference between child self-ratings and parent proxy-ratings of the quality of life of children with a diagnosis of attention-deficit hyperactivity disorder (ADHD)? A systematic review of the literature. *Atten Defic Hyperact Disord, 9*(1), 11-29. doi: 10.1007/s12402-016-0210-9

Garcia-Lopez, L.-J., Saez-Castillo, A. J., & Fuentes-Rodriguez, G. (2017). Psychometric properties of the Youth Anxiety Measure for DSM-5, Part I (YAM-5-I) in a community sample of Spanish-speaking adolescents. *J Affect Disord, 208*, 455-459.

García-Toro, M., Ibarra, O., Gili, M., Serrano, M. J., Oliván, B., Vicens, E., & Roca, M. (2012). Four hygienic-dietary recommendations as add-on treatment in depression: a randomized-controlled trial. *J Affect Disord, 140*(2), 200-203.

Gardner, E., Ruxton, C., & Leeds, A. (2007). Black tea–helpful or harmful? A review of the evidence. *Eur J Clin Nutr, 61*(1), 3.

Garrido, M., Espino, J., González-Gómez, D., Lozano, M., Barriga, C., Paredes, S. D., & Rodríguez, A. B. (2012). The consumption of a Jerte Valley cherry product in humans enhances mood, and increases 5-hydroxyindoleacetic acid but reduces cortisol levels in urine. *Exp Gerontol, 47*(8), 573-580.

Garrison, W. T., & Earls, F. (1985). The child behavior checklist as a screening instrument for young children. *J Am Acad Child Psychiatry, 24*(1), 76-80.

Gelman, R., Baillargeon, R., Mussen, P., Flavell, J., & Markman, E. (1983). Handbook of child psychology: Cognitive development.

Giannakopoulos, G., Tzavara, C., Dimitrakaki, C., Kolaitis, G., Rotsika, V., & Tountas, Y. (2009). The factor structure of the Strengths and Difficulties Questionnaire (SDQ) in Greek adolescents. *Annals of General Psychiatry, 8*(1), 20.

Giles, G. E., Mahoney, C. R., & Kanarek, R. B. (2013). Omega-3 fatty acids influence mood in healthy and depressed individuals. *Nutrition reviews, 71*(11), 727-741.

Gillies, D., Sinn, J., Lad, S. S., Leach, M. J., & Ross, M. J. (2012). Polyunsaturated fatty acids (PUFA) for attention deficit hyperactivity disorder (ADHD) in children and adolescents.

Gohier, B., Ferracci, L., Surguladze, S. A., Lawrence, E., El Hage, W., Kefi, M. Z., . . . Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar depression. *J Affect Disord, 116*(1), 100-105.

Goldacre, B. (2006). The Trial That Ate Itself. *Bad Science.* Retrieved 29/05/2018, 2018, from https://www.badscience.net/2006/09/the-trial-that-ate-itself/

Gomez-Pinilla, F., & Nguyen, T. T. (2012). Natural mood foods: the actions of polyphenols against psychiatric and cognitive disorders. *Nutritional neuroscience, 15*(3), 127-133.

Goodman, Ford, Richards, Gatward, & Meltzer. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(5), 645-655.

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *J Abnorm Child Psychol, 38*(8), 1179-1191. doi: 10.1007/s10802-010-9434-x

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry, 40*(11), 1337-1345. doi: 10.1097/00004583-200111000-00015

Goodman, R., Ford, T., Corbin, T., & Meltzer, H. (2004). Using the Strengths and Difficulties Questionnaire (SDQ) multi-informant algorithm to screen looked-after children for psychiatric disorders. *Eur Child Adolesc Psychiatry, 13 Suppl 2*, Ii25-31. doi: 10.1007/s00787-004-2005-3

Goodman, R., Meltzer, H., & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry, 7*(3), 125-130.

Goodman, R., Meltzer, H., & Bailey, V. (2003). The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Int Rev Psychiatry, 15*(1-2), 173-177. doi: 10.1080/0954026021000046137

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: is small beautiful? *J Abnorm Child Psychol, 27*(1), 17-24.

Children Act (2004).

Guide, S. S. R. U. s. The FACTOR procedure. *Heywood Cases and Other Anomalies about Communality Estimates.* Retrieved 22/06/2018, 2018, from https://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug_factor_sect022.htm

Hansen, A. L., Dahl, L., Olson, G., Thornton, D., Graff, I. E., Frøyland, L., . . . Pallesen, S. (2014). Fish consumption, sleep, daily functioning, and heart rate variability. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine, 10*(5), 567.

Harsha, S., & Anilakumar, K. (2013). Anxiolytic property of Lactuca sativa, effect on anxiety behaviour induced by novel food and height. *Asian Pacific journal of tropical medicine, 6*(7), 532-536.

Hawes, D. J., & Dadds, M. R. (2004). Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *Aust N Z J Psychiatry, 38*(8), 644-651. doi: 10.1080/j.1440-1614.2004.01427.x

Haynes, A., Gilmore, L., Shochet, I., Campbell, M., & Roberts, C. (2013). Factor analysis of the self-report version of the strengths and difficulties questionnaire in a sample of children with intellectual disability. *Res Dev Disabil, 34*(2), 847-854. doi: 10.1016/j.ridd.2012.11.008

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods, 7*(2), 191-205.

Health, D. o. (2010). Liberating the NHS: No Decision About Me, Without Me–Further Consultation on Proposals to Secure Shared Decision-making: Department of Health London.

Heilskov Rytter, M. J., Andersen, L. B. B., Houmann, T., Bilenberg, N., Hvolby, A., Mølgaard, C., . . . Lauritzen, L. (2015). Diet in the treatment of ADHD in children—A systematic review of the literature. *Nord J Psychiatry, 69*(1), 1-18.

Hirose, A., Terauchi, M., Akiyoshi, M., Owa, Y., Kato, K., & Kubota, T. (2016). Low-dose isoflavone aglycone alleviates psychological symptoms of menopause in Japanese women: a randomized, double-blind, placebo-controlled study. *Archives of gynecology and obstetrics, 293*(3), 609-615.

Hirose, A., Terauchi, M., Tamura, M., Akiyoshi, M., Owa, Y., Kato, K., & Kubota, T. (2015). Tomato juice intake increases resting energy expenditure and

improves hypertriglyceridemia in middle-aged women: an open-label, single-arm study. *Nutrition journal, 14*(1), 34.

Hirth, J. M., Rahman, M., & Berenson, A. B. (2011). The association of posttraumatic stress disorder with fast food and soda consumption and unhealthy weight loss behaviors among young women. *Journal of Women's Health, 20*(8), 1141-1149.

Hoyland, Dye, & Lawton. (2009). A systematic review of the effect of breakfast on the cognitive performance of children and adolescents. *Nutrition research reviews, 22*(2), 220-243.

Hoyland, Lawton, & Dye. (2008). Influence of breakfast on cognitive performance, appetite and mood in healthy young adults. *Appetite, 50*(2-3), 560.

Hunt, R. J. (1986). Percent agreement, Pearson's correlation, and kappa as measures of inter-examiner reliability. *Journal of Dental Research, 65*(2), 128-130.

Hurley, L. L., Akinfiresoye, L., Kalejaiye, O., & Tizabi, Y. (2014). Antidepressant effects of resveratrol in an animal model of depression. *Behavioural brain research, 268*, 1-7.

Jaatinen, N., Korpela, R., Poussa, T., Turpeinen, A., Mustonen, S., Merilahti, J., & Peuhkuri, K. (2014). Effects of daily intake of yoghurt enriched with bioactive components on chronic stress responses: a double-blinded randomized controlled trial. *International journal of food sciences and nutrition, 65*(4), 507-514.

Jacka, F. N., Kremer, P. J., Berk, M., de Silva-Sanigorski, A. M., Moodie, M., Leslie, E. R., . . . Swinburn, B. A. (2011). A prospective study of diet quality and mental health in adolescents. *PLoS One, 6*(9), e24805. doi: 10.1371/journal.pone.0024805

Jamieson, S. (2004). Likert scales: how to (ab) use them. *Medical education, 38*(12), 1217-1218.

Janssens, A., & Deboutte, D. (2009). Screening for psychopathology in child welfare: the Strengths and Difficulties Questionnaire (SDQ) compared with the Achenbach System of Empirically Based Assessment (ASEBA). *Eur Child Adolesc Psychiatry, 18*(11), 691-700. doi: 10.1007/s00787-009-0030-y

Jaureguizar, J., Bernaras, E., & Garaigordobil, M. (2017). Child Depression: Prevalence and Comparison between Self-Reports and Teacher Reports. *Spanish Journal of Psychology, 20*. doi: ARTN e17

10.1017/sjp.2017.14

Johnson, S., Hollis, C., Marlow, N., Simms, V., & Wolke, D. (2014). Screening for childhood mental health disorders using the Strengths and Difficulties Questionnaire: the validity of multi-informant reports. *Dev Med Child Neurol, 56*(5), 453-459. doi: 10.1111/dmcn.12360

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement, 20*(1), 141-151.

Kao, G. S., & Thomas, H. M. (2010). Test Review: C. Keith Conners Conners 3rd Edition Toronto, Ontario, Canada: Multi-Health Systems, 2008. *Journal of Psychoeducational Assessment, 28*(6), 598-602.

Kaplan, B. J., Crawford, S. G., Field, C. J., & Simpson, J. S. A. (2007). Vitamins, minerals, and mood. *Psychological Bulletin, 133*(5), 747.

Kaplan, B. J., Crawford, S. G., Gardner, B., & Farrelly, G. (2002). Treatment of mood lability and explosive rage with minerals and vitamins: Two case studies in children. *J Child Adolesc Psychopharmacol, 12*(3), 205-219.

Kaplan, B. J., Fisher, J. E., Crawford, S. G., Field, C. J., & Kolb, B. (2004). Improved mood and behavior during treatment with a mineral-vitamin

supplement: an open-label case series of children. *J Child Adolesc Psychopharmacol, 14*(1), 115-122.

Kean, R. J., Lamport, D. J., Dodd, G. F., Freeman, J. E., Williams, C. M., Ellis, J. A., . . . Spencer, J. P. (2015). Chronic consumption of flavanone-rich orange juice is associated with cognitive benefits: an 8-wk, randomized, double-blind, placebo-controlled trial in healthy older adults–. *Am J Clin Nutr, 101*(3), 506-514.

Keunen, K., Van Elburg, R. M., Van Bel, F., & Benders, M. J. (2014). Impact of nutrition on brain development and its neuroprotective implications following preterm birth. *Pediatric research, 77*(1-2), 148.

Khalid, S., Barfoot, K. L., May, G., Lamport, D. J., Reynolds, S. A., & Williams, C. M. (2017). Effects of acute blueberry flavonoids on mood in children and young adults. *Nutrients, 9*(2), 158.

Khalid, S., Williams, C. M., & Reynolds, S. A. (2016). Is there an association between diet and depression in children and adolescents? A systematic review. *British journal of nutrition, 116*(12), 2097-2108.

Kirby, A., Woodward, A., Jackson, S., Wang, Y., & Crawford, M. (2010). A double-blind, placebo-controlled study investigating the effects of omega-3 supplementation in children aged 8–10 years from a mainstream school population. *Res Dev Disabil, 31*(3), 718-730.

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., . . . Goodman, R. (2000). Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *Eur Child Adolesc Psychiatry, 9*(4), 271-276.

Knowles, J. G., & Cole, A. L. (2008). *Handbook of the arts in qualitative research: Perspectives, methodologies, examples, and issues*: Sage.

Kolenikov, S., & Angeles, G. (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth, 55*(1), 128-165.

Koskelainen, M., Sourander, A., & Kaljonen, A. (2000). The Strengths and Difficulties Questionnaire among Finnish school-aged children and adolescents. *Eur Child Adolesc Psychiatry, 9*(4), 277-284.

Koskelainen, M., Sourander, A., & Vauras, M. (2001). Self-reported strengths and difficulties in a community sample of Finnish adolescents. *Eur Child Adolesc Psychiatry, 10*(3), 180-185.

Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, W. T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry, 160*(9), 1566-1577.

Kresanov, K., Tuominen, J., Piha, J., & Almqvist, F. (1998). Validity of child psychiatric screening methods. *Eur Child Adolesc Psychiatry, 7*(2), 85-95.

Kritz-Silverstein, D., Von Mühlen, D., Barrett-Connor, E., & Bressel, M. A. (2003). Isoflavones and cognitive function in older women: the SOy and Postmenopausal Health In Aging (SOPHIA) Study. *Menopause, 10*(3), 196-202.

Kuhn, C., Aebi, M., Jakobsen, H., Banaschewski, T., Poustka, L., Grimmer, Y., . . . Steinhausen, H. C. (2017). Effective Mental Health Screening in Adolescents: Should We Collect Data from Youth, Parents or Both? *Child Psychiatry Hum Dev, 48*(3), 385-392. doi: 10.1007/s10578-016-0665-0

Kulkarni, S. K., Bhutani, M. K., & Bishnoi, M. (2008). Antidepressant activity of curcumin: involvement of serotonin and dopamine system. *Psychopharmacology (Berl), 201*(3), 435.

Kwan, B., & Rickwood, D. J. (2015). A systematic review of mental health outcome measures for young people aged 12 to 25 years. *BMC Psychiatry, 15*(1), 279.

Lachance, L., & Ramsey, D. (2015). Food, mood, and brain health: Implications for the modern clinician. *Mo Med, 112*(2), 111-115.

Lai, K. Y., Luk, E. S., Leung, P. W., Wong, A. S., Law, L., & Ho, K. (2010). Validation of the Chinese version of the strengths and difficulties questionnaire in Hong Kong. *Soc Psychiatry Psychiatr Epidemiol, 45*(12), 1179-1186. doi: 10.1007/s00127-009-0152-z

Lai, L. S., Hiles, S., Bisquera, A., Hure, A. J., McEvoy, M., & Attia, J. (2013). A systematic review and meta-analysis of dietary patterns and depression in community-dwelling adults–. *Am J Clin Nutr, 99*(1), 181-197.

Lamport, D. J., Lawton, C. L., Merat, N., Jamson, H., Myrissa, K., Hofman, D., . . . Dye, L. (2016). Concord grape juice, cognitive function, and driving performance: a 12-wk, placebo-controlled, randomized crossover trial in mothers of preteen children. *Am J Clin Nutr, 103*(3), 775-783.

Lamport, D. J., Pal, D., Macready, A. L., Barbosa-Boucas, S., Fletcher, J. M., Williams, C. M., . . . Butler, L. T. (2016). The effects of flavanone-rich citrus juice on cognitive function and cerebral blood flow: an acute, randomised, placebo-controlled cross-over trial in healthy, young adults. *British journal of nutrition, 116*(12), 2160-2168.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational research methods, 9*(2), 202-220.

Lansley, A. (2010). Equity and excellence: liberating the NHS. *London: Department of Health*.

Larsson, B., & Frisk, M. (1999). Social competence and emotional/behaviour problems in 6–16 year-old Swedish school children. *Eur Child Adolesc Psychiatry, 8*(1), 24-33.

Law, D., & Wolpert, M. (2014). *Guide to using outcomes and feedback tools with children, young people and families*: CAMHS Press.

le Coq, E. M., Boeke, A. J. P., Bezemer, P. D., Bruil, J., & van Eijk, J. T. M. (2000). Clinimetric properties of a parent report on their offspring's quality of life. *Journal of Clinical Epidemiology, 53*(2), 139-146. doi: Doi 10.1016/S0895-4356(99)00158-4

Le Port, A., Gueguen, A., Kesse-Guyot, E., Melchior, M., Lemogne, C., Nabi, H., . . . Czernichow, S. (2012). Association between dietary patterns and depressive symptoms over time: a 10-year follow-up study of the GAZEL cohort. *PLoS One, 7*(12), e51593.

Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical assessment, research & evaluation, 12*(2), 1-11.

Levy, S. E., & Hyman, S. L. (2003). Use of complementary and alternative treatments for children with autistic spectrum disorders is increasing. *Pediatric Annals, 32*(10), 685-691.

Liu, S. K., Chien, Y. L., Shang, C. Y., Lin, C. H., Liu, Y. C., & Gau, S. S. (2013). Psychometric properties of the Chinese version of Strength and Difficulties Questionnaire. *Compr Psychiatry, 54*(6), 720-730. doi: 10.1016/j.comppsych.2013.01.002

Louis, P. (2012). Does the human gut microbiota contribute to the etiology of autism spectrum disorders? *Digestive Diseases and Sciences*(8), 1987–1989.

Luby, J. L., Belden, A., Sullivan, J., & Spitznagel, E. (2007). Preschoolers' contribution to their diagnosis of depression and anxiety: Uses and limitations of young child self-report of symptoms. *Child Psychiatry & Human Development, 38*(4), 321-338. doi: 10.1007/s10578-007-0063-8

Lundh, L. G., Wangby-Lundh, M., & Bjarehed, J. (2008). Self-reported emotional and behavioral problems in Swedish 14 to 15-year-old adolescents: a study with the self-report version of the Strengths and Difficulties Questionnaire. *Scand J Psychol, 49*(6), 523-532. doi: 10.1111/j.1467-9450.2008.00668.x

Macready, A. L., Kennedy, O. B., Ellis, J. A., Williams, C. M., Spencer, J. P., & Butler, L. T. (2009). Flavonoids and cognitive function: a review of human randomized controlled trial studies and recommendations for future studies. *Genes & nutrition, 4*(4), 227.

Malmberg, M., Rydell, A. M., & Smedje, H. (2003). Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-Swe). *Nord J Psychiatry, 57*(5), 357-363. doi: 10.1080/08039480310002697

Manach, C., & Donovan, J. L. (2004). Pharmacokinetics and metabolism of dietary flavonoids in humans. *Free radical research, 38*(8), 771-785.

Mansbach-Kleinfeld, I., Apter, A., Farbstein, I., Levine, S. Z., & Poznizovsky, A. (2010). A population-based psychometric validation study of the Strengths and Difficulties Questionnaire–Hebrew version. *Frontiers in Psychiatry, 1*, 151.

Marazziti, D., Consoli, G., Picchetti, M., Carlini, M., & Faravelli, L. (2010). Cognitive impairment in major depression. *European journal of pharmacology, 626*(1), 83-86.

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): factor structure, reliability, and validity. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(4), 554-565.

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology, 22*(1), 37.

Martin, K. R., & Appel, C. L. (2009). Polyphenols as dietary supplements: a double-edged sword. *Nutrition and Dietary Supplements, 2*, 1-12.

Mazzio, E. A., Harris, N., & Soliman, K. F. (1998). Food constituents attenuate monoamine oxidase activity and peroxide levels in C6 astrocyte cells. *Planta medica, 64*(07), 603-606.

McDermott, L. M., & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *J Affect Disord, 119*(1), 1-8.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica, 22*(3), 276-282.

McMartin, S. E., Jacka, F. N., & Colman, I. (2013). The association between fruit and vegetable consumption and mental health disorders: evidence from five waves of a national survey of Canadians. *Preventive medicine, 56*(3-4), 225-230.

McQuarrie, A. D., & Tsai, C.-L. (1998). *Regression and time series model selection*: World Scientific.

Mellor, & Moore. (2013). The use of Likert scales with children. *J Pediatr Psychol, 39*(3), 369-379.

Mellor, D. (2004). Furthering the use of the strengths and difficulties questionnaire: reliability with younger child respondents. *Psychol Assess, 16*(4), 396-401. doi: 10.1037/1040-3590.16.4.396

Mellor, D. (2005). Normative data for the Strengths and Difficulties Questionnaire in Australia. *Australian psychologist, 40*(3), 215-222.

Mellor, D., & Stokes, M. (2007). The factor structure of the Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 23*(2), 105.

Mellor, D., Wong, J., & Xu, X. (2011). Interparent agreement on the strengths and difficulties questionnaire: A Chinese study. *Journal of Clinical Child & Adolescent Psychology, 40*(6), 890-896.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2000). *The mental health of children and adolescents in Great Britain*: HM Stationery Office.

Mhurchu, C. N., Gorton, D., Turley, M., Jiang, Y., Michie, J., Maddison, R., & Hattie, J. (2013). Effects of a free school breakfast programme on children's attendance, academic achievement and short-term hunger: results from a stepped-wedge, cluster randomised controlled trial. *Journal of Epidemiology & Community Health, 67*(3), 257-264.

Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and reliability of the strengths and difficulties questionnaire in 5-6 year olds: differences by gender or by parental education? *PLoS One, 7*(5), e36805. doi: 10.1371/journal.pone.0036805

Mihrshahi, S., Dobson, A., & Mishra, G. (2015). Fruit and vegetable consumption and prevalence and incidence of depressive symptoms in mid-age women: results from the Australian longitudinal study on women's health. *Eur J Clin Nutr, 69*(5), 585.

Miller, M. G., & Shukitt-Hale, B. (2012). Berry fruit enhances beneficial signaling in the brain. *Journal of agricultural and food chemistry, 60*(23), 5709-5715.

Mulle, J. G., Sharp, W. G., & Cubells, J. F. (2013). The gut microbiome: a new frontier in autism research. *Current psychiatry reports, 15*(2), 337.

Muratori, P., Giuli, C., Bertacchi, I., Orsolini, L., Ruglioni, L., & Lochman, J. E. (2017). Coping power for preschool-aged children: a pilot randomized control trial study. *Early Interv Psychiatry, 11*(6), 532-538.

Muris, P., Meesters, C., de Kanter, E., & Timmerman, P. E. (2005). Behavioural inhibition and behavioural activation system scales for children: relationships with Eysenck's personality traits and psychopathological symptoms. *Personality and Individual Differences, 38*(4), 831-841.

Muris, P., Meesters, C., & Diederen, R. (2005). Psychometric properties of the Big Five Questionnaire for Children (BFQ-C) in a Dutch sample of young adolescents. *Personality and Individual Differences, 38*(8), 1757-1769.

Muris, P., Meesters, C., Eijkelenboom, A., & Vincken, M. (2004). The self-report version of the Strengths and Difficulties Questionnaire: its psychometric properties in 8- to 13-year-old non-clinical children. *Br J Clin Psychol, 43*(Pt 4), 437-448. doi: 10.1348/0144665042388982

Muris, P., Meesters, C., & Schouten, E. (2002). A brief questionnaire of DSM-IV-defined anxiety and depression symptoms among children. *Clinical Psychology & Psychotherapy, 9*(6), 430-442.

Muris, P., Meesters, C., & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ)--further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *Eur Child Adolesc Psychiatry, 12*(1), 1-8. doi: 10.1007/s00787-003-0298-2

Muris, P., Merckelbach, H., Schmidt, H., & Mayer, B. (1998). The revised version of the Screen for Child Anxiety Related Emotional Disorders (SCARED-R): Factor structure in normal children. *Personality and Individual Differences, 26*(1), 99-112.

Muris, P., Simon, E., Lijphart, H., Bos, A., Hale, W., & Schmeitz, K. (2017). The Youth Anxiety Measure for DSM-5 (YAM-5): development and first psychometric evidence of a new scale for assessing anxiety disorders symptoms of children and adolescents. *Child Psychiatry & Human Development, 48*(1), 1-17.

Murphy, Moore, Tapper, Lynch, Clarke, Raisanen, . . . Moore. (2011). Free healthy breakfasts in primary schools: a cluster randomised controlled trial of a policy intervention in Wales, UK. *Public health nutrition, 14*(2), 219-226.

Murphy, J. M., Pagano, M. E., Nachmani, J., Sperling, P., Kane, S., & Kleinman, R. E. (1998). The relationship of school breakfast to psychosocial and academic functioning: cross-sectional and longitudinal observations in an inner-city school sample. *Arch Pediatr Adolesc Med, 152*(9), 899-907.

Muthayya, S., Thomas, T., Srinivasan, K., Rao, K., Kurpad, A. V., van Klinken, J.-W., . . . de Bruin, E. A. (2007). Consumption of a mid-morning snack improves memory but not attention in school children. *Physiology & behavior, 90*(1), 142-150.

Myers, K., & Winters, N. C. (2002a). Ten-Year Review of Rating Scales. I. *J Am Acad Child Adolesc Psychiatry, 41*(2), 114-122.

Myers, K., & Winters, N. C. (2002b). Ten-year review of rating scales. I: overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(2), 114-122.

Naismith, S. L., Hickie, I. B., Turner, K., Little, C. L., Winter, V., Ward, P. B., . . . Parker, G. (2003). Neuropsychological performance in patients with depression is associated with clinical, etiological and genetic risk factors. *Journal of clinical and experimental neuropsychology, 25*(6), 866-877.

Nemets, B., Stahl, Z., & Belmaker, R. (2002). Addition of omega-3 fatty acid to maintenance medication treatment for recurrent unipolar depressive disorder. *American Journal of Psychiatry, 159*(3), 477-479.

Niditch, L. A., & Varela, R. E. (2011). Mother-Child Disagreement in Reports of Child Anxiety: Effects of Child Age and Maternal Anxiety. *J Anxiety Disord, 25*(3), 450-455. doi: 10.1016/j.janxdis.2010.11.009

Norbury, C. F., Gooch, D., Baird, G., Charman, T., Simonoff, E., & Pickles, A. (2016). Younger children experience lower levels of language competence and academic progress in the first year of school: evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(1), 65-73.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education, 15*(5), 625-632.

O'Neil, A., Quirk, S. E., Housden, S., Brennan, S. L., Williams, L. J., Pasco, J. A., . . . Jacka, F. N. (2014). Relationship between diet and mental health in children and adolescents: a systematic review. *American journal of public health, 104*(10), e31-e42.

Oddy, W. H., Robinson, M., Ambrosini, G. L., Therese, A., de Klerk, N. H., Beilin, L. J., . . . Stanley, F. J. (2009). The association between dietary patterns and mental health in early adolescence. *Preventive medicine, 49*(1), 39-44.

Opie, R. S., O'Neil, A., Itsiopoulos, C., & Jacka, F. N. (2015). The impact of whole-of-diet interventions on depression and anxiety: a systematic review of randomised controlled trials. *Public health nutrition, 18*(11), 2074-2093.

Palmieri, P. A., & Smith, G. C. (2007). Examining the structural validity of the Strengths and Difficulties Questionnaire (SDQ) in a US sample of custodial grandmothers. *Psychol Assess, 19*(2), 189.

Pase, M. P., Scholey, A. B., Pipingas, A., Kras, M., Nolidin, K., Gibbs, A., . . . Stough, C. (2013). Cocoa polyphenols enhance positive mood states but not cognitive performance: a randomized, placebo-controlled trial. *Journal of psychopharmacology, 27*(5), 451-458.

Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A., & Ploner, A. (2005). False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics, 21*(13), 3017-3024.

Peet, M., & Horrobin, D. F. (2002). A dose-ranging study of the effects of ethyl-eicosapentaenoate in patients with ongoing depression despite apparently adequate treatment with standard drugs. *Arch Gen Psychiatry, 59*(10), 913-919.

Percy, A., McCrystal, P., & Higgins, K. (2008). Confirmatory factor analysis of the adolescent self-report Strengths and Difficulties Questionnaire. *European Journal of Psychological Assessment, 24*(1), 43-48.

Piaget, J. (1954). Language and thought from a genetic perspective. *Acta Psychologica, 10*, 51-60.

Pinheiro, L. C., McFatrich, M., Lucas, N., Walker, J. S., Withycombe, J. S., Hinds, P. S., . . . Mack, J. W. (2017). Child and adolescent self-report symptom measurement in pediatric oncology research: a systematic literature review. *Quality of Life Research*, 1-29.

Piqueras, J. A., Kuhne, W., Vera-Villarroel, P., Van Straten, A., & Cuijpers, P. (2011). Happiness and health behaviours in Chilean college students: a cross-sectional survey. *BMC Public Health, 11*(1), 443.

Piqueras, J. A., Martín-Vivar, M., Sandin, B., San Luis, C., & Pineda, D. (2017). The Revised Child Anxiety and Depression Scale: A systematic review and reliability generalization meta-analysis. *J Affect Disord, 218*, 153-169. doi: 10.1016/j.jad.2017.04.022

Pivik, R., Tennal, K. B., Chapman, S. D., & Gu, Y. (2012). Eating breakfast enhances the efficiency of neural networks engaged during mental arithmetic in school-aged children. *Physiology & behavior, 106*(4), 548-555.

Psaltopoulou, T., Sergentanis, T. N., Panagiotakos, D. B., Sergentanis, I. N., Kosti, R., & Scarmeas, N. (2013). Mediterranean diet, stroke, cognitive impairment, and depression: a meta-analysis. *Annals of neurology, 74*(4), 580-591.

Puri, B. K., Counsell, S. J., Richardson, A. J., & Horrobin, D. F. (2002). Eicosapentaenoic acid in treatment-resistant depression. *Arch Gen Psychiatry, 59*(1), 91-92.

Raine, A., Portnoy, J., Liu, J., Mahoomed, T., & Hibbeln, J. R. (2015). Reduction in behavior problems with omega-3 supplementation in children aged 8–16 years: a randomized, double-blind, placebo-controlled, stratified, parallel-group trial. *Journal of Child Psychology and Psychiatry, 56*(5), 509-520.

Reeve, B. B., McFatrich, M., Pinheiro, L. C., Weaver, M. S., Sung, L., Withycombe, J. S., . . . Gibson, D. (2017). Eliciting the child's voice in adverse event reporting in oncology trials: Cognitive interview findings from the Pediatric Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events initiative. *Pediatric blood & cancer, 64*(3).

Reichelt, A. C., Westbrook, R. F., & Morris, M. J. (2017). Impact of Diet on Learning, Memory and Cognition. *Frontiers in behavioral neuroscience, 11*, 96.

Rendeiro, C., Vauzour, D., Rattray, M., Waffo-Téguo, P., Mérillon, J. M., Butler, L. T., . . . Spencer, J. P. (2013). Dietary levels of pure flavonoids improve spatial memory performance and increase hippocampal brain-derived neurotrophic factor. *PLoS One, 8*(5), e63535.

Rescorla, L. A., Bochicchio, L., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., . . . Verhulst, F. C. (2014). Parent-teacher agreement on children's problems in 21 societies. *J Clin Child Adolesc Psychol, 43*(4), 627-642. doi: 10.1080/15374416.2014.900719

Richter, L. M., Rose, C., & Griesel, R. D. (1997). Cognitive and behavioural effects of a school breakfast. *South African Medical Journal, 87*(1).

Rienks, J., Dobson, A., & Mishra, G. (2013). Mediterranean dietary pattern and prevalence and incidence of depressive symptoms in mid-aged women: results from a large community-based prospective study. *Eur J Clin Nutr, 67*(1), 75.

Rønning, J. A., Handegaard, B. H., Sourander, A., & Mørch, W.-T. (2004). The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *Eur Child Adolesc Psychiatry, 13*(2), 73-82.

Ronning, J. A., Handegaard, B. H., Sourander, A., & Morch, W. T. (2004). The Strengths and Difficulties Self-Report Questionnaire as a screening instrument in Norwegian community samples. *Eur Child Adolesc Psychiatry, 13*(2), 73-82. doi: 10.1007/s00787-004-0356-4

Ruchkin, V., Jones, S., Vermeiren, R., & Schwab-Stone, M. (2008). The strengths and difficulties questionnaire: The self-report version in American urban and suburban youth. *Psychol Assess, 20*(2), 175.

Rucklidge, Johnstone, & Kaplan. (2009). Nutrient supplementation approaches in the treatment of ADHD. *Expert review of neurotherapeutics, 9*(4), 461-476.

Rucklidge, J., Taylor, M., & Whitehead, K. (2011). Effect of micronutrients on behavior and mood in adults with ADHD: Evidence from an 8-week open label trial with natural extension. *J Atten Disord, 15*(1), 79-91.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *Journal of Child Psychology and Psychiatry, 8*(1), 1-11.

Ruusunen, A., Lehto, S. M., Mursu, J., Tolmunen, T., Tuomainen, T.-P., Kauhanen, J., & Voutilainen, S. (2014). Dietary patterns are associated with the prevalence of elevated depressive symptoms and the risk of getting a hospital discharge diagnosis of depression in middle-aged or older Finnish men. *J Affect Disord, 159*, 1-6.

Sánchez-Villegas, A., Delgado-Rodríguez, M., Alonso, A., Schlatter, J., Lahortiga, F., Majem, L. S., & Martínez-González, M. A. (2009). Association of the Mediterranean dietary pattern with the incidence of depression: the Seguimiento Universidad de Navarra/University of Navarra follow-up (SUN) cohort. *Arch Gen Psychiatry, 66*(10), 1090-1098.

Sánchez-Villegas, A., Martínez-González, M. A., Estruch, R., Salas-Salvadó, J., Corella, D., Covas, M. I., . . . Lapetra, J. (2013). Mediterranean dietary pattern and depression: the PREDIMED randomized trial. *BMC medicine, 11*(1), 208.

Sánchez-Villegas, A., Verberne, L., De Irala, J., Ruíz-Canela, M., Toledo, E., Serra-Majem, L., & Martínez-González, M. A. (2011). Dietary fat intake and the risk of depression: the SUN Project. *PLoS One, 6*(1), e16268.

Sanhueza, C., Ryan, L., & Foxcroft, D. (2013). Diet and the risk of unipolar depression in adults: systematic review of cohort studies. *Journal of Human Nutrition and Dietetics, 26*(1), 56-70.

Sanne, B., Torsheim, T., Heiervang, E., & Stormark, K. M. (2009). The Strengths and Difficulties Questionnaire in the Bergen Child Study: a conceptually and methodically motivated structural analysis. *Psychol Assess, 21*(3), 352.

Sathyapalan, T., Beckett, S., Rigby, A. S., Mellor, D. D., & Atkin, S. L. (2010). High cocoa polyphenol rich chocolate may reduce the burden of the symptoms in chronic fatigue syndrome. *Nutrition journal, 9*(1), 55.

Sawyer, M. G., Baghurst, P., & Mathias, J. (1992). Differences Between Informants' Reports describing Emotional and Behavioural Problems in Community and Clinic-referred Children: a Research Note. *Journal of Child Psychology and Psychiatry, 33*(2), 441-449.

Sayegh, R., Schiff, I., Wurtman, J., Spiers, P., McDERMOTT, J., & Wurtman, R. (1995). The effect of a carbohydrate-rich beverage on mood, appetite, and cognitive function in women with premenstrual syndrome. *Obstetrics & Gynecology, 86*(4), 520-528.

Scalbert, A., & Williamson, G. (2000). Dietary intake and bioavailability of polyphenols. *J Nutr, 130*(8), 2073S-2085S.

Scholey, A., & Owen, L. (2013). Effects of chocolate on cognitive function and mood: a systematic review. *Nutrition reviews, 71*(10), 665-681.

Scholey, A. B., French, S. J., Morris, P. J., Kennedy, D. O., Milne, A. L., & Haskell, C. F. (2010). Consumption of cocoa flavanols results in acute improvements in mood and cognitive performance during sustained mental effort. *Journal of psychopharmacology, 24*(10), 1505-1514.

Schwab-Stone, M., Fallon, T., Briggs, M., & Crowther, B. (1994). Reliability of diagnostic reporting for children aged 6–11 years: A test-retest study of the Diagnostic Interview Schedule for Children—Revised. *The American journal of psychiatry*.

Selma, M. V., Espin, J. C., & Tomas-Barberan, F. A. (2009). Interaction between phenolics and gut microbiota: role in human health. *Journal of agricultural and food chemistry, 57*(15), 6485-6501.

Silverman, W. K., & Eisen, A. R. (1992). Age differences in the reliability of parent and child reports of child anxious symptomatology using a structured interview. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*(1), 117-124.

Skarupski, K. A., Tangney, C., Li, H., Evans, D., & Morris, M. (2013). Mediterranean diet and depressive symptoms among older adults over time. *The journal of nutrition, health & aging, 17*(5), 441-445.

Slobodskaya, H. R. (2007). The associations among the big five, behavioural inhibition and behavioural approach systems and child and adolescent adjustment in Russia. *Personality and Individual Differences, 43*(4), 913-924.

Smedje, H., Broman, J.-E., Hetta, J., & von Knorring, A.-L. (1999). Psychometric properties of a Swedish version of the "Strengths and Difficulties Questionnaire". *Eur Child Adolesc Psychiatry, 8*(2), 63-70. doi: 10.1007/s007870050086

Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., . . . Döpfner, M. (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.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing behavior problems. *Journal of Clinical Child Psychology, 22*(1), 107-116.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical assessment, research & evaluation, 9*(4), 1-19.

Stoll, A. L., Severus, W. E., Freeman, M. P., Rueter, S., Zboyan, H. A., Diamond, E., . . . Marangell, L. B. (1999). Omega 3 fatty acids in bipolar disorder: a preliminary double-blind, placebo-controlled trial. *Arch Gen Psychiatry, 56*(5), 407-412.

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clin Child Fam Psychol Rev, 13*(3), 254-274. doi: 10.1007/s10567-010-0071-2

Streiner, D. L. (2003). Being inconsistent about consistency: When coefficient alpha does and doesn't matter. *J Pers Assess, 80*(3), 217-222.

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., . . . Greenhill, L. L. (2001). Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 168-179.

Szliszka, E., & Krol, W. (2011). The role of dietary polyphenols in tumor necrosis factor-related apoptosis inducing ligand (TRAIL)-induced apoptosis for cancer chemoprevention. *European Journal of Cancer Prevention, 20*(1), 63-69.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*: Allyn & Bacon/Pearson Education.

Taib, M. N. M., Shariff, Z. M., Wesnes, K. A., Saad, H. A., & Sariman, S. (2012). The effect of high lactose–isomaltulose on cognitive performance of young children. A double blind cross-over design study. *Appetite, 58*(1), 81-87.

Tobias, S., & Carlson, J. E. (1969). Brief report: Bartlett's test of sphericity and chance findings in factor analysis. *Multivariate Behavioral Research, 4*(3), 375-377.

Trebatická, J., & Ďuračková, Z. (2015). Psychiatric disorders and polyphenols: can they be helpful in therapy? *Oxidative medicine and cellular longevity, 2015*.

Trebatická, J., Kopasová, S., Hradečná, Z., Činovský, K., Škodáček, I., Šuba, J., . . . Rohdewald, P. (2006). Treatment of ADHD with French maritime pine bark extract, Pycnogenol®. *Eur Child Adolesc Psychiatry, 15*(6), 329-335.

Tuohy, K. M., Conterno, L., Gasperotti, M., & Viola, R. (2012). Up-regulating the human intestinal microbiome using whole plant foods, polyphenols, and/or fiber. *Journal of agricultural and food chemistry, 60*(36), 8776-8782.

Ullebø, A. K., Posserud, M.-B., Heiervang, E., Obel, C., & Gillberg, C. (2012). Prevalence of the ADHD phenotype in 7-to 9-year-old children: effects of informant, gender and non-participation. *Soc Psychiatry Psychiatr Epidemiol, 47*(5), 763-769.

UNICEF. (1989). The United Nations Convention on the Rights of the Child. London: UNICEF UK.

Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2006). Self-reported strengths and difficulties in a large Norwegian population 10-19 years : age and gender specific results of the extended SDQ-questionnaire. *Eur Child Adolesc Psychiatry, 15*(4), 189-198. doi: 10.1007/s00787-005-0521-4

Van Roy, B., Groholt, B., Heyerdahl, S., & Clench-Aas, J. (2010). Understanding discrepancies in parent-child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. *BMC Psychiatry, 10*(1), 56. doi: 10.1186/1471-244x-10-56

Van Roy, B., Veenstra, M., & Clench-Aas, J. (2008). Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. *J Child Psychol Psychiatry, 49*(12), 1304-1312. doi: 10.1111/j.1469-7610.2008.01942.x

van Widenfelt, B. M., Goedhart, A. W., Treffers, P. D., & Goodman, R. (2003). Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *Eur Child Adolesc Psychiatry, 12*(6), 281-289. doi: 10.1007/s00787-003-0341-3

Vauzour, D., Vafeiadou, K., Rodriguez-Mateos, A., Rendeiro, C., & Spencer, J. P. (2008). The neuroprotective potential of flavonoids: a multiplicity of effects. *Genes & nutrition, 3*(3), 115.

Vaz, S., Cordier, R., Boyes, M., Parsons, R., Joosten, A., Ciccarelli, M., . . . Falkmer, T. (2016). Is using the strengths and difficulties questionnaire in a community sample the optimal way to assess mental health functioning? *PLoS One, 11*(1), e0144039.

Verhulst, F. C., & Ende, J. (1992). Agreement between parents' reports and adolescents' self-reports of problem behavior. *Journal of Child Psychology and Psychiatry, 33*(6), 1011-1023.

Vinot, N., Jouin, M., Lhomme-Duchadeuil, A., Guesnet, P., Alessandri, J.-M., Aujard, F., & Pifferi, F. (2011). Omega-3 fatty acids from fish oil lower anxiety, improve cognitive functions and reduce spontaneous locomotor activity in a non-human primate. *PLoS One, 6*(6), e20491.

Von Baeyer, C. L., Carlson, G., & Webb, L. (1997). Underprediction of pain in children undergoing ear piercing. *Behaviour research and therapy, 35*(5), 399-404.

Wardle, J., Rogers, P., Judd, P., Taylor, M. A., Rapoport, L., Green, M., & Nicholson Perry, K. (2000). Randomized trial of the effects of cholesterol-lowering dietary treatment on psychological function. *Am J Med, 108*(7), 547-553.

Wasowski, C., & Marder, M. (2012). Flavonoids as GABAA receptor ligands: the whole story? *Journal of experimental pharmacology, 4*, 9.

Watson, A. W., Haskell-Ramsay, C. F., Kennedy, D. O., Cooney, J. M., Trower, T., & Scheepens, A. (2015). Acute supplementation with blackcurrant extracts modulates cognitive functioning and inhibits monoamine oxidase-B in healthy young adults. *Journal of Functional Foods, 17*, 524-539.

Weems, G. H., & Onwuegbuzie, A. J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement and Evaluation in Counseling and Development, 34*(3), 166.

Weis, R., & Smenner, L. (2007). Construct validity of the Behavior Assessment System for Children (BASC) Self-Report of Personality: Evidence from adolescents referred to residential treatment. *Journal of Psychoeducational Assessment, 25*(2), 111-126.

Wender, E. H., & Solanto, M. V. (1991). Effects of sugar on aggressive and inattentive behavior in children with attention deficit disorder with hyperactivity and normal children. *Pediatrics, 88*(5), 960-966.

Wesnes, K. A., Pincock, C., Richardson, D., Helm, G., & Hails, S. (2003). Breakfast reduces declines in attention and memory over the morning in schoolchildren. *Appetite, 41*(3), 329-331.

Wesnes, K. A., Pincock, C., & Scholey, A. (2012). Breakfast is associated with enhanced cognitive function in schoolchildren. An internet based study. *Appetite, 59*(3), 646-649.

White, B. A., Horwath, C. C., & Conner, T. S. (2013). Many apples a day keep the blues away–Daily experiences of negative and positive affect and food consumption in young adults. *British Journal of Health Psychology, 18*(4), 782-798.

Whyte, A. R., Schafer, G., & Williams, C. M. (2016). Cognitive effects following acute wild blueberry supplementation in 7-to 10-year-old children. *European journal of nutrition, 55*(6), 2151-2162.

Whyte, A. R., & Williams, C. M. (2015). Effects of a single dose of a flavonoid-rich blueberry drink on memory in 8 to 10 y old children. *Nutrition, 31*(3), 531-534.

Widenhorn-Müller, K., Hille, K., Klenk, J., & Weiland, U. (2008). Influence of having breakfast on cognitive performance and mood in 13-to 20-year-old high school students: results of a crossover trial. *Pediatrics, 122*(2), 279-284.

Williams, L. J., Brennan, S. L., Henry, M. J., Berk, M., Jacka, F. N., Nicholson, G. C., . . . Pasco, J. A. (2011). Area-based socioeconomic status and mood disorders: cross-sectional evidence from a cohort of randomly selected adult women. *Maturitas, 69*(2), 173-178.

Woerner, Becker, & Rothenberger. (2004). Normative data and scale properties of the German parent SDQ. *Eur Child Adolesc Psychiatry, 13*(2), ii3-ii10. doi: 10.1007/s00787-004-2002-6

Woerner, W., Becker, A., Friedrich, C., Klasen, H., Goodman, R., & Rothenberger, A. (2002). [Normal values and evaluation of the German parents' version of Strengths and DIfficulties Questionnaire (SDQ): Results of a representative field study]. *Z Kinder Jugendpsychiatr Psychother, 30*(2), 105-112. doi: 10.1024//1422-4917.30.2.105

Wolpert, M., Deighton, J., De Francesco, D., Martin, P., Fonagy, P., & Ford, T. (2014). From 'reckless' to 'mindful'in the use of outcome data to inform service-level performance management: perspectives from child mental health. *BMJ Qual Saf*, bmjqs-2013-002557.

Wolpert, M., Ford, T., Trustam, E., Law, D., Deighton, J., Flannery, H., & Fugard, R. J. (2012). Patient-reported outcomes in child and adolescent mental health services (CAMHS): Use of idiographic and standardized measures. *Journal of Mental Health, 21*(2), 165-173.

Wolpert, M., Fugard, A. J., Deighton, J., & Görzig, A. (2012). Routine outcomes monitoring as part of children and young people's Improving Access to Psychological Therapies (CYP IAPT)–improving care or unhelpful burden? *Child and Adolescent Mental Health, 17*(3), 129-130.

Yao, S., Zhang, C., Zhu, X., Jing, X., McWhinnie, C. M., & Abela, J. R. (2009). Measuring adolescent psychopathology: psychometric properties of the self-report strengths and difficulties questionnaire in a sample of Chinese adolescents. *J Adolesc Health, 45*(1), 55-62. doi: 10.1016/j.jadohealth.2008.11.006

Youdim, M. B., & Bakhle, Y. (2006). Monoamine oxidase: isoforms and inhibitors in Parkinson's disease and depressive illness. *British journal of pharmacology, 147*(S1).

Zhang, J.-c., Yao, W., Dong, C., Yang, C., Ren, Q., Ma, M., . . . Hashimoto, K. (2015). Comparison of ketamine, 7, 8-dihydroxyflavone, and ANA-12 antidepressant effects in the social defeat stress model of depression. *Psychopharmacology (Berl), 232*(23), 4325-4335.

Zuccato, C., & Cattaneo, E. (2009). Brain-derived neurotrophic factor in neurodegenerative diseases. *Nature Reviews Neurology, 5*(6), 311.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of modern applied statistical methods, 6*(1), 4.

Appendices

Appendix 1: Questionnaire Structure

Questionnaire	Informant	Score	Calculation	Range
SDQ (25-item)	All	Emotional problems sub-scale	Sum of items: 3, 8, 13, 16, 24	0 - 10
		Conduct problems sub-scale	Sum of items: 5, 7, 12, 18, 22	0 - 10
		Hyperactivity sub-scale	Sum of items: 2, 10, 15, 21, 25	0 - 10
		Peer problems sub-scale	Sum of items: 6, 11, 14, 19, 23	0 - 10
		Prosocial Strengths sub-scale	Sum of items: 1, 4, 9, 17, 20	0 - 10
		Total Difficulty composite score	Sum of all scales, except the prosocial scale	0 - 40
		Externalising Difficulties composite score	Sum of conduct and hyperactivity scales	0 - 20
		Internalising Difficulties composite score	Sum of emotional and peer problems scales	0 – 20
RCADS (47-item)	Parent	Separation Anxiety (SA) sub-scale	Sum of items: 5, 9, 17, 18, 33, 45, 46	0 - 21
		Social Phobia (SP) sub-scale	Sum of items: 4, 7, 8, 12, 20, 30, 32, 38, 43	0 - 27
		Generalised anxiety Disorder (GAD) sub-scale	Sum of items: 1, 13, 22, 27, 35, 37	0 - 18
		Panic Disorder (PD) sub-scale	Sum of items: 3, 14, 24, 26, 28, 34, 36, 39, 41	0 - 27
		Obsessive compulsive disorder (OCD) sub-scale	Sum of items: 10, 16, 23, 31, 42, 44	0 - 18
		Major depressive disorder (MDD) sub-scale	Sum of items: 2, 6, 11, 15, 19, 21, 25, 29, 40, 47	0 - 30
		Total composite score	Sum of all sub-scales	0 - 141

RCADS (25-item)	Self-report	Separation Anxiety (SA) sub-scale	Sum of items: 3, 6, 9	0 - 9
		Social Phobia (SP) sub-scale	Sum of items: 2, 7, 22	0 - 9
		Generalised anxiety Disorder (GAD) sub-scale	Sum of items: 5, 18, 25	0 - 9
		Panic Disorder (PD) sub-scale	Sum of items: 11, 14, 20	0 - 9
		Obsessive compulsive disorder (OCD) sub-scale	Sum of items: 12, 17, 23	0 - 9
		Total Depression sub-scale	Sum of items: 1, 4, 8, 10, 13, 15, 16, 19, 21, 24	0 - 30
		Total Anxiety composite score	Sum of all sub-scales (except Total Depression)	0 - 45
		Total composite score	Sum of all sub-scales	0 – 75
SNAP-IV (18-item)	Parent and Teacher	Inattention	Sum of items: 1 – 9 divided by number of items (9)	0 – 3
		Hyperactivity / Impulsivity	Sum of items: 10 – 18 divided by number of items (9)	0 - 3
		Total	Sum of items: 1 – 18 divided by the number of items (18)	0 - 3

Appendix 2: Strengths and Difficulties Questionnaire Self-Report Version

	Not True	Somewhat True	Certainly True
I try to be nice to other people. I care about their feelings	☐	☐	☐
I am restless, I cannot stay still for long	☐	☐	☐
I get a lot of headaches, stomach-aches or sickness	☐	☐	☐
I usually share with others (food, games, pens etc.)	☐	☐	☐
I get very angry and often lose my temper	☐	☐	☐
I am usually on my own. I generally play alone or keep to myself	☐	☐	☐
I usually do as I am told	☐	☐	☐
I worry a lot	☐	☐	☐
I am helpful if someone is hurt, upset or feeling ill	☐	☐	☐
I am constantly fidgeting or squirming	☐	☐	☐
I have one good friend or more	☐	☐	☐
I fight a lot. I can make other people do what I want	☐	☐	☐
I am often unhappy, down-hearted or tearful	☐	☐	☐
Other people my age generally like me	☐	☐	☐
I am easily distracted, I find it difficult to concentrate	☐	☐	☐
I am nervous in new situations. I easily lose confidence	☐	☐	☐
I am kind to younger children	☐	☐	☐
I am often accused of lying or cheating	☐	☐	☐
Other children or young people pick on me or bully me	☐	☐	☐
I often volunteer to help others (parents, teachers, children)	☐	☐	☐
I think before I do things	☐	☐	☐
I take things that are not mine from home, school or elsewhere	☐	☐	☐
I get on better with adults than with people my own age	☐	☐	☐
I have many fears, I am easily scared	☐	☐	☐
I finish the work I'm doing. My attention is good	☐	☐	☐

Do you have any other comments or concerns?

Appendix 3: Strengths and Difficulties Questionnaire Informant Version

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

Do you have any other comments or concerns?

Appendix 4: Revised Child Anxiety and Depression Scale (RCADS), 25-Item, Self-Report Version

Name: _____ Date: _____

RCADS – Short Version

Please put a circle around the word that shows how often each of these things happen to you. There are no right or wrong answers.

1. I feel sad or empty	Never	Sometimes	Often	Always
2. I worry when I think I have done poorly at something	Never	Sometimes	Often	Always
3. I would feel afraid of being on my own at home	Never	Sometimes	Often	Always
4. Nothing is much fun anymore	Never	Sometimes	Often	Always
5. I worry that something awful will happen to someone in my family	Never	Sometimes	Often	Always
6. I am afraid of being in crowded places (like shopping centers, the movies, buses, busy playgrounds)	Never	Sometimes	Often	Always
7. I worry what other people think of me	Never	Sometimes	Often	Always
8. I have trouble sleeping	Never	Sometimes	Often	Always
9. I feel scared if I have to sleep on my own	Never	Sometimes	Often	Always
10. I have problems with my appetite	Never	Sometimes	Often	Always
11. I suddenly become dizzy or faint when there is no reason for this	Never	Sometimes	Often	Always
12. I have to do some things over and over again (like washing my hands, cleaning or putting things in a certain order)	Never	Sometimes	Often	Always
13. I have no energy for things	Never	Sometimes	Often	Always
14. I suddenly start to tremble or shake when there is no reason for this	Never	Sometimes	Often	Always
15. I cannot think clearly	Never	Sometimes	Often	Always
16. I feel worthless	Never	Sometimes	Often	Always
17. I have to think of special thoughts (like numbers or words) to stop bad things from happening	Never	Sometimes	Often	Always
18. I think about death	Never	Sometimes	Often	Always
19. I feel like I don't want to move	Never	Sometimes	Often	Always
20. I worry that I will suddenly get a scared feeling when there is nothing to be afraid of	Never	Sometimes	Often	Always
21. I am tired a lot	Never	Sometimes	Often	Always
22. I feel afraid that I will make a fool of myself in front of people	Never	Sometimes	Often	Always
23. I have to do some things in just the right way to stop bad things from happening	Never	Sometimes	Often	Always
24. I feel restless	Never	Sometimes	Often	Always
25. I worry that something bad will happen to me	Never	Sometimes	Often	Always

Appendix 5: Revised Child Anxiety and Depression Scale (RCADS), 47-Item Parent Version

Date: _____ **RCADS-P**
Name/ID: _____ Relationship to Child: _____

Please put a circle around the word that shows how often each of these things happens for your child.

1. My child worries about things	Never	Sometimes	Often	Always
2. My child feels sad or empty	Never	Sometimes	Often	Always
3. When my child has a problem, he/she gets a funny feeling in his/her stomach	Never	Sometimes	Often	Always
4. My child worries when he/she thinks he/she has done poorly at something	Never	Sometimes	Often	Always
5. My child feels afraid of being alone at home	Never	Sometimes	Often	Always
6. Nothing is much fun for my child anymore	Never	Sometimes	Often	Always
7. My child feels scared when taking a test	Never	Sometimes	Often	Always
8. My child worries when he/she thinks someone is angry with him/her.	Never	Sometimes	Often	Always
9. My child worries about being away from me	Never	Sometimes	Often	Always
10. My child is bothered by bad or silly thoughts or pictures in his/her mind	Never	Sometimes	Often	Always
11. My child has trouble sleeping	Never	Sometimes	Often	Always
12. My child worries about doing badly at school work	Never	Sometimes	Often	Always
13. My child worries that something awful will happen to someone in the family	Never	Sometimes	Often	Always
14. My child suddenly feels as if he/she can't breathe when there is no reason for this.	Never	Sometimes	Often	Always
15. My child has problems with his/her appetite	Never	Sometimes	Often	Always
16. My child has to keep checking that he/she has done things right (like the switch is off, or the door is locked)	Never	Sometimes	Often	Always
17. My child feels scared to sleep on his/her own	Never	Sometimes	Often	Always
18. My child has trouble going to school in the mornings because of feeling nervous or afraid.	Never	Sometimes	Often	Always
19. My child has no energy for things	Never	Sometimes	Often	Always
20. My child worries about looking foolish	Never	Sometimes	Often	Always
21. My child is tired a lot	Never	Sometimes	Often	Always
22. My child worries that bad things will happen to him/her	Never	Sometimes	Often	Always
23. My child can't seem to get bad or silly thoughts out of his/her head.	Never	Sometimes	Often	Always

24. When my child has a problem, his/her heart beats really fast	Never	Sometimes	Often	Always
25. My child cannot think clearly	Never	Sometimes	Often	Always
26. My child suddenly starts to tremble or shake when there is no reason for this	Never	Sometimes	Often	Always
27. My child worries that something bad will happen to him/her	Never	Sometimes	Often	Always
28. When My child has a problem, he/she feels shaky	Never	Sometimes	Often	Always
29. My child feels worthless	Never	Sometimes	Often	Always
30. My child worries about making mistakes	Never	Sometimes	Often	Always
31. My child has to think of special thoughts (like numbers or words) to stop bad things from happening	Never	Sometimes	Often	Always
32. My child worries what other people think of him/her	Never	Sometimes	Often	Always
33. My child is afraid of being in crowded places (like shopping centers, the movies, buses, busy playgrounds)	Never	Sometimes	Often	Always
34. All of a sudden my child will feel really scared for no reason at all	Never	Sometimes	Often	Always
35. My child worries about what is going to happen	Never	Sometimes	Often	Always
36. My child suddenly becomes dizzy or faint when there is no reason for this	Never	Sometimes	Often	Always
37. My child thinks about death	Never	Sometimes	Often	Always
38. My child feels afraid if he/she have to talk in front of the class	Never	Sometimes	Often	Always
39. My child's heart suddenly starts to beat too quickly for no reason	Never	Sometimes	Often	Always
40. My child feels like he/she doesn't want to move	Never	Sometimes	Often	Always
41. My child worries that he/she will suddenly get a scared feeling when there is nothing to be afraid of	Never	Sometimes	Often	Always
42. My child has to do some things over and over again (like washing hands, cleaning, or putting things in a certain order)	Never	Sometimes	Often	Always
43. My child feels afraid that he/she will make a fool of him/herself in front of people	Never	Sometimes	Often	Always
44. My child has to do some things in just the right way to stop bad things from happening	Never	Sometimes	Often	Always
45. My child worries when in bed at night	Never	Sometimes	Often	Always
46. My child would feel scared if he/she had to stay away from home overnight	Never	Sometimes	Often	Always
47. My child feels restless	Never	Sometimes	Often	Always

Appendix 6: Swanson, Nolan and Pelham – IV Questionnaire (SNAP-IV)

SNAP-IV Teacher and Parent 18-Item Rating Scale

James M. Swanson, Ph.D., University of California, Irvine, CA 92715

Patient/Client Name: _____

Date of birth: _____ Gender: _____

Grade: _____ Type of class: _____ Class size: _____

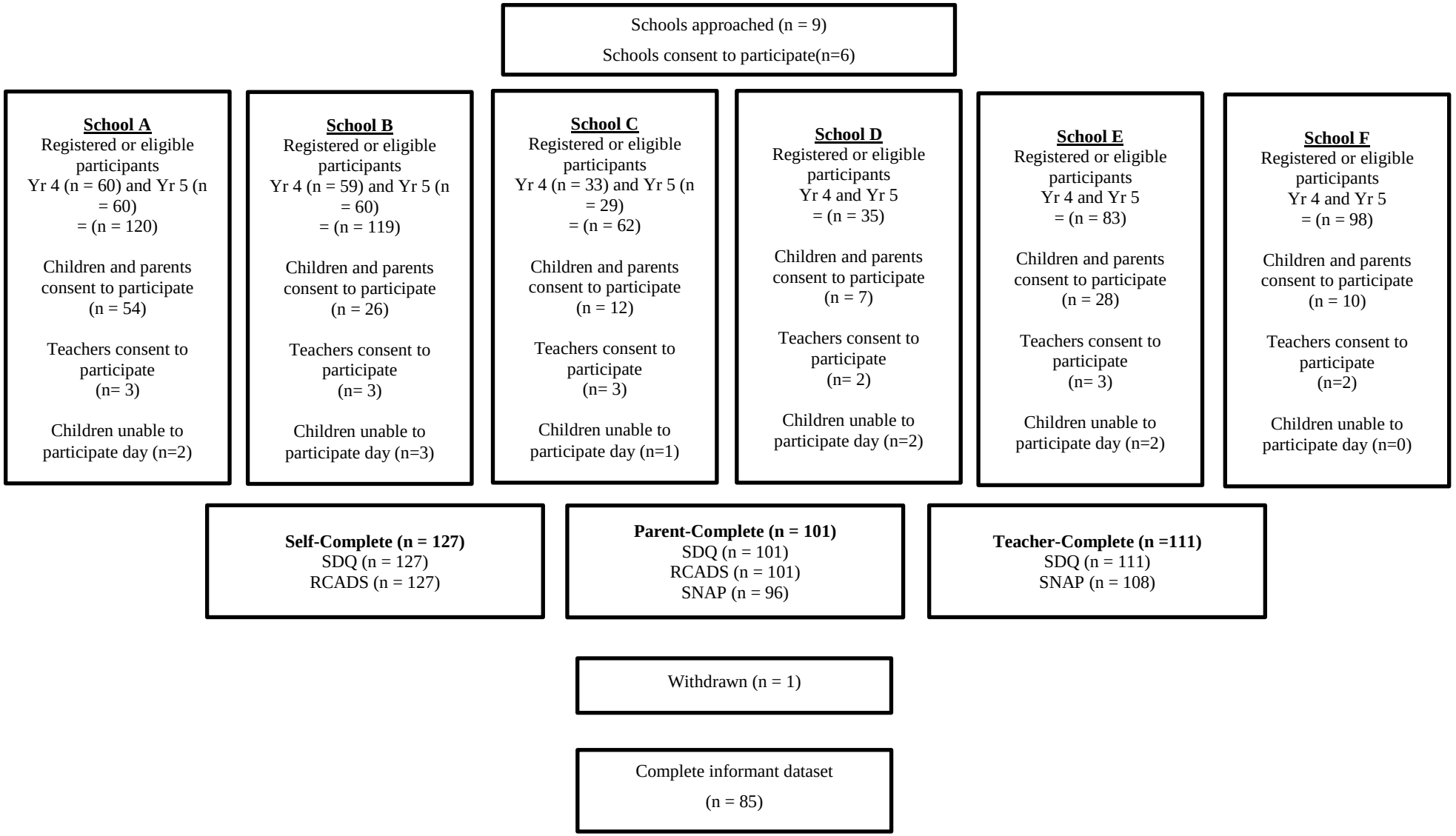
Completed by: _____ Date: _____

Physician Name: _____

For each item, check the column which best describes this child/adolescent:

	Not at all	Just a little	Quite a bit	Very much
1. Often fails to give close attention to details or makes careless mistakes in schoolwork or tasks				
2. Often has difficulty sustaining attention in tasks or play activities				
3. Often does not seem to listen when spoken to directly				
4. Often does not follow through on instructions and fails to finish schoolwork, chores, or duties				
5. Often has difficulty organizing tasks and activities				
6. Often avoids, dislikes, or reluctantly engages in tasks requiring sustained mental effort				
7. Often loses things necessary for activities (e.g., toys, school assignments, pencils or books				
8. Often is distracted by extraneous stimuli				
9. Often is forgetful in daily activities				
10. Often fidgets with hands or feet or squirms in seat				
11. Often leaves seat in classroom or in other situations in which remaining seated is expected				
12. Often runs about or climbs excessively in situations in which it is inappropriate				
13. Often has difficulty playing or engaging in leisure activities quietly				
14. Often is "on the go" or often acts as if "driven by a motor"				
15. Often talks excessively				
16. Often blurts out answers before questions have been completed				
17. Often has difficulty awaiting turn				
18. Often interrupts or intrudes on others (e.g., butts into conversations/ games				

Appendix 7: Flow Diagram of Recruitment for Tool Development



Appendix 8: Tool Development - Parent information sheet (and opt-in consent form)

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

Dear Parent/Carer,

A study to assess the self-report of behaviour and well-being in 8-10 year old school children

My name is Dr Fiona Wright. I am a clinical psychologist in training from the University of Leeds. As part of my thesis I am conducting a study comparing the views of the child, the parent and the teacher of behaviour and emotional difficulties of young children, through the use of questionnaires. This letter invites you and your child to participate.

Normally, when we consider behaviour and emotional difficulties in young children (under the age of 12) we ask the adults. But, it makes sense that the children themselves may be better at telling us what is going on inside of them. Disappointingly, most of the questionnaires available that allow children to tell us about these issues themselves, have been tested only in older children and adolescents. So this study is hoping to show that a younger child's self-report of behaviour and emotional difficulties are just as valuable.

We would like to invite you and your child to complete some questionnaires about your child's wellbeing and behaviour. If you consent, your child will complete two questionnaires and you will complete three questionnaires. With your consent we will also ask your child's teacher to complete two questionnaires about your child's wellbeing and behaviour. We expect that the parent questionnaires should take no more than 30 minutes to complete. These will be made available in online or paper format.

This letter is designed to provide you with enough information about the study and the questionnaires in order for you to make an informed decision about you and your child's participation. If you are happy for your child to participate, we will then also discuss the questionnaires with your child using an accessible (child friendly) information sheet and ask if they are happy to take part.

What the study entails:

We would ask your child and your child's teacher to complete a total of two questionnaires each and you to complete a total of three questionnaires (all of which ask questions about your child's behaviour and well-being).

The benefits of taking part in this research:

We will enter your child into a prize draw, as a thank you for participating in the study. This gives your child the chance to win one £40, two £20, and one £10 Love2Shop voucher.

The Questionnaires

There will be three questionnaires to assess your child's behaviour and well-being:

The Strengths and Difficulties Questionnaire (SDQ)

The SDQ assesses children's emotional and behavioural strengths and difficulties. As part of the study, this will be completed in a self-report format by your child at school, with assistance from me. It will also be completed by their teacher and by yourself.

As their parent/guardian you are ideally placed to give us day-to-day knowledge about your child. The SDQ requires you to read a series of statements and to judge how well each statement describes your child by ticking one of three or four boxes for each question. Completion shouldn't take long – about 5 to 10 minutes on average.

The Revised Child Anxiety and Depression Scale (RCADS)

The RCADS assesses signs of anxiety and depression in children. As part of the study, this will be completed in a short self-report format by your child at school, with assistance from me. A slightly longer version will also be completed by yourself. Completion should take about 10 – 15 minutes on average. Again, this requires you to read a series of statements and judge how well each statement describes your child by circling one of four options.

The SNAP-4

The SNAP-4 assesses attention and activity levels in children. As part of the study, this will be completed by your child's teacher and by yourself. This questionnaire requires that you read a series of statements and judge how well each statement describes your child by ticking one of four options. This questionnaire should take about 5 – 10 minutes to complete.

Support

It is important that you understand the questions asked within the questionnaires. If you have any problems completing the parent questionnaires please contact me (see below). If

necessary I can arrange a home visit to assist you in completing the forms, or I can answer any questions or talk it through over the phone.

It is also important to remember that it is expected that your child will be in the normal range for these questionnaire statements. However, if any of the questions raise concerns for you please contact me (see below).

Further Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 17-0253; date of approval 02/10/2017). All of the information collected from you, your child and their teacher during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If you consent for you and your child to take part, you are both free to withdraw at any time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if you and your child participate in the study, you can still request the withdrawal of your and their information up until completion of the research (May 2018). If you wish to remove you and your child's data from the study at a later date please contact me (details below) or let your child's teacher know. At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. The results from the study will be used towards an educational qualification and published in international scientific journals.

This study will be ongoing in school during the period September – December 2017. All researchers will provide full DBS clearance prior to undertaking the study.

If you are you happy for your child to participate in this study then please inform the school using the slip below. If you do not wish for your child to participate you do not need to do anything further.

If you have any questions about this study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Please reply by **.**

I will also be at the school on *** between ***** and I would be happy to discuss any questions about the sub-study.**

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

A study to assess the self-report of behaviour and well-being in 8-10 year old school children

I **DO** wish my son/daughter .....

(Please insert child's name)

to take part in the above research study. I am also happy to complete three questionnaires about my child's behaviour and well-being and for my child's teacher to be asked to complete two questionnaires about my child's behaviour and well-being.

Signed.....(parent/carer)

Date.....

Contact Telephone Number .....

Email Address: ……

Please only return this slip if you DO wish your child to take part in this sub-study.

Please reply by **

Appendix 9: Tool Development - Screening Questionnaire

School of Psychology

Leeds LS2 9JT

PLEASE COMPLETE AND RETURN IF

YOU **DO** WISH YOUR CHILD TO TAKE PART

Name of child……..

1. Which **school** does your child attend?

☐ School #1

☐ School #2

☐ School #3

☐ School #4

☐ School #5

2. Which **Year Group** is your child in?

☐ Year 4

☐ Year 5

3. What is your child's date of birth?

_____ / _____ / _____

4. Which **gender** is your child?

☐ Male

☐ Female

5. Does your child take any **medication**?

☐ Yes

☐ No

If yes, please state _____

Please reply by XX/XX/XXXX at the latest

Now please return your questionnaire back to School

Appendix 10: Tool Development - Parent information sheet (to be sent alongside questionnaires)

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

Dear Parent/Carer,

A study to assess the self-report of behaviour and well-being in 8-10 year old school children

Thank you for already consenting for your child to take part in this study.

This letter should give you information about filling in the parent versions of the behaviour and well-being questionnaires. We ask that you complete these questionnaires within the next two weeks (by XX/XX/XXXX at the latest)

Remember: We will enter your child into a prize draw, as thanks for participating in the study. This gives your child the chance to win one £40, two £20, and one £10 Love2Shop voucher.

Your child has already completed two questionnaires asking about their well-being and mood. The two questionnaires are called (1) The **Strengths and Difficulties Questionnaire** and (2) The **Revised Child Anxiety and Depression Scale**. Your child filled these out themselves at school with my help. The first one is looking at your child's strengths and difficulties with feelings and behaviour. The second is looking worries and low mood.

We now ask you to fill in three questionnaires to see your views about your child's behaviour and well-being. These include (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale, in addition to (3) the **SNAP-IV**. The additional questionnaire asks about your child's attention and activity levels.

To access the questionnaires please visit: **http/ LINK**

Alternatively, please contact me (details below) for paper copies.

As their parent/guardian you are the best person to give us day-to-day knowledge about your child. All the questionnaires ask you to read a series of sentences and judge how well it describes your child by ticking one of the options. It shouldn't take long.

Support

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me (see below). If necessary, I can arrange a home visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that your child will be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please contact me (see below).

Contact Details

If you have any questions about this study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Alternatively you can email or phone the project supervisors at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Please complete the questionnaires by XX/XX/XXXX at the latest

Appendix 11: Tool Development - Parent information (as seen prior to the online completion of the questionnaires – Bristol Online Survey)

A study to assess the self-report of behaviour and well-being in 8-10 year old school children

Thank you for already consenting for your child to take part in this study.

Remember: We will enter your child into a prize draw, as thanks for participating in the study. This gives your child the chance to win Love2Shop vouchers.

Your child has already completed two questionnaires asking about their well-being and mood. The two questionnaires are called (1) The **Strengths and Difficulties Questionnaire** and (2) The **Revised Child Anxiety and Depression Scale**. Your child filled these out themselves at school with my help. The first one is looking at your child's strengths and difficulties with feelings and behaviour. The second is looking worries and low mood.

We now ask you to fill in three questionnaires to see your views about your child's behaviour and well-being. These include (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale, in addition to (3) the **SNAP-IV**. The additional questionnaire is thinking about your child's attention and activity levels.

As their parent/guardian you are the best person to give us day-to-day knowledge about your child. All the questionnaires ask you to read a series of sentences and judge how well it describes your child by ticking one of the options. It shouldn't take long.

If you would rather complete these questionnaires in a paper format, please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

<u>Support</u>

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me. If necessary, I can arrange a home visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that your child will be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please contact me.

PAGE 2

Ethical Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 17-0253; date of approval 02/10/2017)

Before continuing with the questionnaire, I confirm that I have previously read and understood the information sheet provided (V2 28/09/2017) and signed the consent form. I understand that my participation is voluntary and that I am free to withdraw at any time until completion of the research (May 2018) without giving any reason. I understand that all information that is collected from me, my child and their teacher during the course of the study will be kept confidential and will only be used for the purposes of this research. I agree take part in this study and complete the following questionnaires.

YES

NO

Appendix 12: Tool Development - Parent information (as seen post the online completion of the questionnaires – Bristol Online Survey)

Thank you for completing this questionnaire

If you are successful in winning the Love2Shop vouchers in the prize draw, you will be contacted before 31st January 2018

If you have any further questions about this study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Alternatively you can email or phone the project supervisors at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Appendix 13: Tool Development - Teacher information sheet (pre-study)

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

Dear Teachers and Teaching Assistants,

A study assessing the self-report of behaviour and well-being in 8-10 year old UK school children

My name is Dr Fiona Wright. I am a clinical psychologist in training from the University of Leeds. As part of my thesis I am conducting a study looking into the role of self-reported measures of behaviour and well-being in your school.

If the parent/guardian's consent, children and their parents will complete some questionnaires looking at the behaviour and well-being of the child. As part of the study, we are asking that the children's teachers also complete two questionnaires relating to the child's behaviour and well-being.

This letter is designed to provide you with information about the study and the behaviour and well-being questionnaires to be used in this study.

What the study entails:

We would ask you (as the child's teacher) to complete a total of two questionnaires for each participating child in your class. The child will also complete two questionnaires, with support from myself during school hours. Their parents/guardians will complete three questionnaires (all of which ask questions about their child's behaviour and well-being).

The benefits of taking part in this research:

As a thank you for taking part in the study you will be given £15 in Love2Shop vouchers at the end of the study.

NB: Alternatively, the school may have chosen to use this money to provide teaching cover (to allow you time to complete the questionnaires).

The Questionnaires

There will be three measures to assess behaviour and well-being:

The Strengths and Difficulties Questionnaire (SDQ)

The SDQ is about assessing children's emotional and behavioural strengths and difficulties. Looking at emotions, conduct, attention, peer relationships and prosocial behaviour. It exists in a number of versions for child self-report, parent and teachers.

As their teacher, the SDQ asks you to read a series of statements and judge how well it describes your pupil by ticking one fo the three boxes for each question. This should take about 5-10 minutes on average.

The SNAP-4

The SNAP-4 is about assessing the children's attention and activity levels. As their teacher, the SNAP-4 again requires you to read a series of statements and judge how well it described your child by ticking one of four options. Completion should take 5 – 10 minutes, on average.

and

The Revised Child Anxiety and Depression Scale (RCADS)

The RCADS is about assessing signs of anxiety and depression. This will be completed in a short self-report format by the children at school, with assistance from the research team. A slightly longer version will also be completed by the children's' parents/guardians. Teachers will not be asked to complete this questionnaire as there is no teacher version.

Support

It is important that you understand the questions asked within the questionnaires, if you have any problems completing the forms please contact me (see below).

It is also important to remember that it is expected that your pupil will be in the low to normal range for these statements. However, if this does raise concerns for you please contact me (see below).

Further Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 17-0253; date of approval 02/10/2017). All of the information collected from the children, parents and teachers during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If the children decide to take part they are free to withdraw at any time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if a child participates in the study, they can still request the withdrawal of their information up until completion of work for publication (May 2018). If a child wishes to remove their data from the study at a later date please contact one of the research team (details below). At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. The results from the study will be used towards an educational qualification and published in international scientific journals.

This study will be ongoing in school during the period September – December 2017. All researchers will provide full DBS clearance prior to undertaking the study.

If you have any questions about this part of the study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Appendix 14: Tool Development - Teacher information sheet (to be sent alongside questionnaires)

School of Psychology

Leeds LS2 9JT

Dear Teachers and Teaching Assistants,

A study assessing the self-report of behaviour and well-being in 8-10 year old UK school children

My name is Dr Fiona Wright. I am a clinical psychologist in training from the University of Leeds. As part of my thesis I am conducting a study looking into the role of self-reported measures of behaviour and well-being in your school.

This letter gives you information about filling in the teacher versions of the behaviour and well-being questionnaires. We ask that you complete these questionnaires within the next two weeks (by XX/XX/XXXX at the latest)

Remember: As a thank you for taking part in the study you will be given £15 in Love2Shop vouchers at the end of the study. NB: Alternatively, the school may have chosen to use this money to provide teaching cover (to allow you time to complete the questionnaires).

A number of parent/guardians have consented to their child taking part in this study. These children have already completed two questionnaires asking about their well-being and mood. The two questionnaires are called (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale. They filled these out themselves at school with my help.

As part of the study, we are asking you (as their teacher) to complete two questionnaires, for each of the children in your class who are taking part in the study. These include (1) The **Strengths and Difficulties Questionnaire** and (2) the **SNAP-IV**. The first one is about

assessing your pupil's emotional and behavioural strengths and difficulties. The second is about assessing your pupil's attention and activity levels.

To access the questionnaires please visit: **http/ LINK**

Alternatively, please contact me (details below) for paper copies.

As their teacher you are ideally placed to give us day-to-day knowledge about your pupils. All the questionnaires ask you to read a series of statements and judge how well it describes your pupil by ticking one of the options. It shouldn't take long – no more than 10 minutes.

Support:

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me (see below). If necessary I can arrange a school visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that your pupils will all be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please

contact me (see below).

Contact Details

If you have any questions about this part of the study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Alternatively you can email or phone the project supervisors at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Please complete the questionnaires by XX/XX/XXXX at the latest

Appendix 15: Tool Development - Teacher information (as seen prior to the online completion of the questionnaires – Bristol Online Survey)

A study to assess the self-report of behaviour and well-being in 8-10 year old school children

Thank you for participating in this study. The child who you are completing these questionnaires for has already consented to participate in the study.

Remember: As a thank you for taking part in the study you will be given £15 in Love2Shop vouchers at the end of the study. NB: Alternatively, the school may have chosen to use this money to provide teaching cover (to allow you time to complete the questionnaires).

We are asking you to fill in two questionnaires about this child to see your views about their behaviour and well-being before. These include (1) The Strengths and Difficulties Questionnaire and (2) the **SNAP-IV**.

As their teacher you are ideally placed to give us day-to-day knowledge about your pupils. All the questionnaires ask you to read a series of statements and judge how well it describes your pupil by ticking one of the options. It shouldn't take long – no more than 10 minutes.

If you would rather complete these questionnaires in a paper format, please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Support

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me. If necessary, I can arrange a school visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that all the children in your class will be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please contact me.

Ethical Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 17-0253; date of approval 02/10/2017)

Before continuing with the questionnaire, I confirm that I have previously read and understood the information sheet provided (V2 28/09/2017). I understand that my participation is voluntary and that I am free to withdraw at any time until completion of the research (May 2018) without giving any reason. I understand that all information that is collected from me during the course of the study will be kept confidential and will only be used for the purposes of this research. I agree take part in this study and complete the following questionnaires.

Yes

No

Appendix 16: Tool Development - Teacher information (as seen post the online completion of the questionnaires – Bristol Online Survey)

Thank you for completing this questionnaire

The Love2Shop vouchers will be distributed before 31st January 2018

If you have any further questions about this study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

Alternatively you can email or phone the project supervisors at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Appendix 17: Tool Development - Child Information and Consent form

UNIVERSITY OF LEEDS

Hi there!

Thank you for agreeing to take part in this study

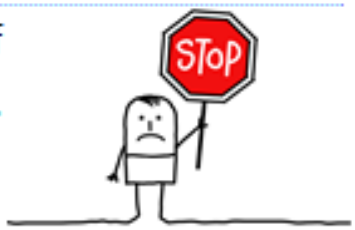
We would like your help to find out whether your view about your behaviour and mood matches that of your parents and teachers.

What you will need to do:

- You will fill in two paper questionnaires

If there are any questions that you don't want to answer, you don't have to— you can skip these

You do not have to take part if you don't want to. That is okay. Just tell us or your teacher



If you do take part, you can change your mind at any time and stop without telling us why. Just tell us or your teacher if you want to stop.

We will use your results to write reports for other researchers. Nobody will know your name, only a number that is linked to your results so no one can tell the information came from you.

If you have any questions or if you are unsure about anything, please ask us.

You do NOT have to decide RIGHT NOW if you want to take part. Take your time to think about it.

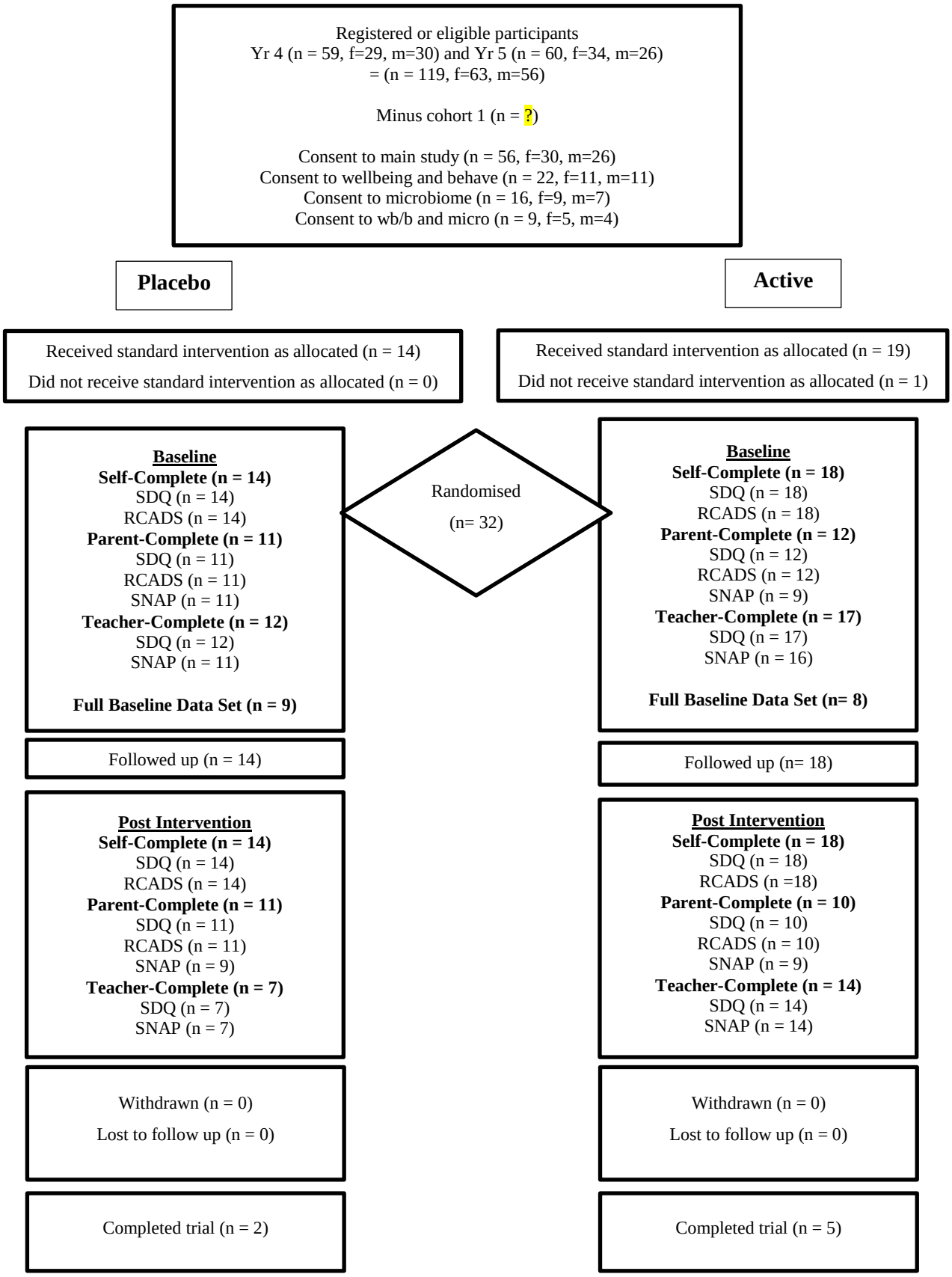
If you do not want to take part, that is okay. You can go back to class

If you do want to take part we can start

If you understand what the study is about and are happy to continue, please sign your name below

_____

Appendix 18: CONSORT-based Flow Diagram of Recruitment for the Thesis Intervention Study



Appendix 19: Thesis Intervention Study, Cohort 1 - Behaviour and Well-being Sub-study Parent Information Sheet (and opt-in consent)

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

Dear Parent/Carer,

An intervention study of the effects of 100% fresh orange juice on behaviour and well-being in 7-10 year old school children

Professor Louise Dye from the University of Leeds is currently leading a team who are conducting a study in your child's school. My name is Dr Fiona Wright, also from the University of Leeds.

In addition to the ongoing study examining how different types of orange-based drinks affect children's cognition, we would like you to consent to your child completing some additional questionnaires. We would like to examine how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks, in addition to the cognitive measures we are already taking at school. There is some evidence to suggest that healthy foods can affect well-being and behaviour of children in a positive way.

If you consent, your child will complete two additional questionnaires and three questionnaires will be sent home for yourselves to complete. We expect that this should take no more than 30 minutes at two time point.

This letter is designed to provide you with enough information about the study in order for you to make an informed decision about your child's participation. If you are happy for your child to participate, we will then also discuss the study with each child using an accessible information sheet and ask if they are happy to take part.

What the study entails:

Following the initial screening of the primary study, we would ask you, your child and your child's teacher to complete a total of three questionnaires to examine your child's behaviour and well-being prior to receiving one of the orange drinks.

Following the 10 week orange juice intervention, we would again ask you, your child and your child's teacher to complete these questionnaires, to examine the impact of the orange-based drinks on their behaviour and well-being.

The Questionnaires

There will be three additional measures to assess behaviour and well-being:

The Strengths and Difficulties Questionnaire (SDQ)

The SDQ is about assessing your child's emotional and behavioural strengths and difficulties. As part of the study, this will be completed in a self-report format by your child at school, with assistance. It will also be completed by their teacher and yourself.

As their parent/guardian you are ideally placed to give us day-to-day knowledge about your child. The SDQ requires you to read a series of statements and judge how well it describes your child by ticking one of three boxes for each question. Completion shouldn't take long – about 5 to 10 minutes on average.

The Revised Child Anxiety and Depression Scale (RCADS)

The RCADS is about assessing signs of anxiety and depression in your child. As part of the study, this will be completed in a short self-report format by you child at school, with assistance. A slightly longer version will also be completed by yourself. Completion should take about 10 – 15 minutes on average. Again, this requires you to read a series of statements and judge how well it describes your child by circling one of four options.

and

The SNAP-4

The SNAP-4 is about assessing your child's attention and activity levels. As part of the study, this will be completed by your child's teacher and yourself. This questionnaire requires that you read a series of statements and judge how well it describes your child by ticking one of four options, and should take about 5 – 10 minutes to complete.

Support

It is important that you understand the questions asked within the questionnaires, if you have any problems completing the forms please contact me (see below). If necessary I can arrange a home visit to assist you in completing the forms, or I can answer any questions or talk it through over the phone.

It is important to remember that it is expected that your child will be in the low to normal range for these statements. However, if this does raise concerns for you please contact me (see below).

Further Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (16-0181 approved 29/06/2016). All of the information collected from your child during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If you decide to allow your child to take part they are free to withdraw at any time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if your child participates in the study, you can still request the withdrawal of their information up until completion of work for publication (July 2017). If you wish to remove your child's data from the study at a later date please contact one of the research team (details below) or let your child's teacher know. At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. Some results from the study will be used towards an educational qualification by members of the research team and published in international scientific journals.

This study will be ongoing in school during the period February –December 2016. All researchers will provide full DBS clearance prior to undertaking the study.

If you are you happy for your child to participate in this study then please inform the school using the slip below. If you do not wish for your child to participate you do not need to do anything further.

If you have any questions about this part of the study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

If you have any questions about the main study please contact:

Fiona Croden on 0113 343 5753 or Neil Boyle 0113 343 0540 email: n.b.boyle@leeds.ac.uk

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

An intervention study of the effects of 100% fresh orange juice on behaviour and well-being in 7-10 year old school children

I **DO** wish my son/daughter .....

(Please insert child's name)

to take part in the above research study.

Signed.....(parent/carer)

Date.....

Contact number .....

Please only return this slip if you DO wish your child to take part in this study.

Please reply by __/__/__

Appendix 20: Thesis Intervention Study, Cohort 1 - Behaviour and Well-being Sub-study Child Information Sheet/Consent

UNIVERSITY OF LEEDS

Hi there!

Thank you for agreeing to take part in the orange juice study.

We would like your help to find out whether the drink you are getting has an effect on your behaviour and mood.

What you will need to do:

Have orange juice every morning for 10 weeks

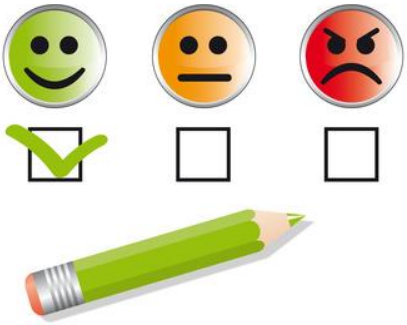
In week 1, **before** you start drinking the orange juice:

- You will fill in two paper questionnaires

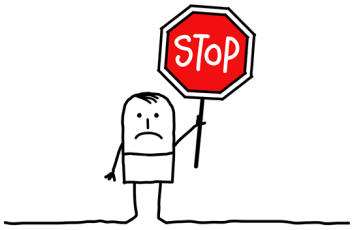
After 10 weeks of drinking the juice:

- You will fill in the same paper questionnaires again!

If there are any questions that you don't want to answer, you don't have to— you can skip these



You do not have to take part if you don't want to. That is okay. Just tell us or your teacher



If you do take part, you can change your mind at any time and stop without telling us why. Just tell us or your teacher if you want to stop.

We will use your results to write reports for other researchers. Nobody will know your name, only a number that is linked to your results so no one can tell the information came from you.

If you have any questions or if you are unsure about anything, please ask us.

If you do not want to take part, that is okay. You can go back to class

If you do want to take part we can start

If you understand what the study is about and are happy to continue, please sign your name below

Appendix 21: Thesis Intervention Study, Cohort 1 - Behaviour and Well-being Teacher Information Sheet

School of Psychology

Leeds LS2 9JT

Dear Teachers and Teaching Assistants,

An intervention study of the effects of 100% fresh orange juice on behaviour and well-being in 7-10 year old school children

Professor Louise Dye from the University of Leeds is currently leading a team who are conducting a study in your pupil's school. My name is Dr Fiona Wright, also from the University of Leeds. In addition to the ongoing study examining how different types of orange-based drinks affect children's cognition, we would like to examine how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks.

If the parent/guardian's consent, these children and their parents will complete some questionnaires, in addition to the cognitive measures we are already taking at school. There is some evidence to suggest that healthy foods can affect well-being and behaviour of children in a positive way.

As part of the study, we are asking that the children's teachers complete two questionnaires relating to the child's behaviour and well-being across two time points.

This letter is designed to provide you with information about the additional behaviour and well-being measures to be used in this study.

What the study entails:

Following the initial screening of the primary study, we would ask you, your pupil and your pupil's parent/guardian to complete a total of three of questionnaires to examine their behaviour and well-being prior to receiving one of the orange drinks.

Following the 10 week orange juice intervention, we would again ask you, your pupil and your pupil's parent/guardian to complete these questionnaires again, to examine the impact of the orange-based drinks on their behaviour and well-being.

The Questionnaires

There will be three additional measures to assess behaviour and well-being:

• **The Strengths and Difficulties Questionnaire (SDQ)**
The SDQ is about assessing children's emotional and behavioural strengths and difficulties. Looking at emotions, conduct, attention, peer relationships and prosocial behaviour. It exists in a number of versions for child self-report, parent and teachers.

As their teacher, the SDQ asks you to read a series of statements and judge how well it describes your pupil by ticking one fo the three boxes for each question. This should take about 5-10 mintues on average.

• **The Revised Child Anxiety and Depression Scale (RCADS)**

The RCADS is about assessing signs of anxiety and depression. This will be completed in a short self-report format by the children at school, with assistance from the research team. A slightly longer version will also be completed by the children's' parents/guardians.

and

• **The SNAP-4**

The SNAP-4 is about assessing the children's attention and activity levels. As their teacher, the SNAP-4 again requires you to read a series of statements and judge how well it described your child by ticking one of four options. Completion should take 5 – 10 minutes, on average.

Support

It is important that you understand the questions asked within the questionnaires, if you have any problems completing the forms please contact me (see below).

It is also important to remember that it is expected that your pupil will be in the low to normal range for these statements. However, if this does raise concerns for you please contact me (see below).

Further Details

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 16-0181 approved 29/06/2016). All of the information collected from the children during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If the children decide to take part they are free to withdraw at any time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if a child participates in the study, they can still request the withdrawal of their information up until completion of work for publication (July 2017). If a child wishes to remove their data from the study at a later date please contact one of the research team (details below). At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. Some results from the study will be used towards an educational qualification by members of the research team and published in international scientific journals.

This study will be ongoing in school during the period February –December 2016. All researchers will provide full DBS clearance prior to undertaking the study.

If you have any questions about this part of the study please contact me:

Fiona Wright on 07956024807 or ps07fw@leeds.ac.uk

If you have any questions about the main study please contact:

Fiona Croden on 0113 343 5753 or Neil Boyle 0113 343 0540 email: n.b.boyle@leeds.ac.uk

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Appendix 22: Thesis Intervention Study, Cohort 2 - Combined Parent Informed Sheet (and opt-in consent form and screen questionnaire for all study elements combined)

School of Psychology
Leeds LS2 9JT

Dear Parent/Carer,

An intervention study of the effects of 100% fresh orange juice on

(1) cognitive performance
(2) behaviour and well-being
and

(3) gut bacteria
in 8-10 year old school children

My name is Professor Louise Dye from the University of Leeds. I am leading a team who are conducting a study in your child's school. This work is a research collaboration between the University of Leeds and the State of Florida Department of Citrus.

We would like to invite your child to take part in a study examining how different types of orange-based drinks affect children's performance and well-being over a period of 10 weeks. The study drinks are A) 100% fresh pure orange juice B) an orange based drink containing the same amount of vitamins/fructose.

This letter is designed to provide you with enough information about the study in order for you to make an informed decision about your child's participation. If you are happy for your child to participate, we will then also discuss the study with each child using an accessible information sheet and ask if they are happy to take part.

What the study entails:

At an initial screening visit we will administer a short IQ and colour blindness test. If your child is found to be colour-blind we will pass this information on to you via the school. They will also complete a peak flow test to measure lung capacity. This is to examine if activity levels affect cognitive performance.

Finally, your child will complete two questionnaires asking about their well-being and mood. The two questionnaires are called (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale. These will be completed in a self-report format by your child at school, with assistance. The first one is about assessing your child's emotional and behavioural strengths and difficulties. The second is about assessing signs of anxiety and depression in your child. We would like to examine how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks as there is some evidence to suggest that healthy foods can affect the well-being and behaviour of children in a positive way.

Following the initial screening, we would ask you and your child's teacher to complete a total of three questionnaires to examine your child's behaviour and well-being prior to receiving one of the orange drinks. These will include (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale, in addition to (3) the SNAP-IV. The additional questionnaire is about assessing your child's attention and activity levels. As their parent/guardian you are ideally placed to give us day-to-day knowledge about your child. The questionnaires require you to read a series of statements and judge how well it describes your child by ticking one of the options. Completion shouldn't take long – about 15 minutes on average

For the gut bacteria bit, we would also ask that that your child provides a small faecal (stool) sample at home before they start receiving one of the orange-based drinks. We will provide you with a specially designed stool collection kit, instructions on how to use this kit and freepost postal boxes (one for each sample) so that you can post the samples off for analysis. . This analysis will be carried out by our collaborator (Dr David Vauzour from the University of East Anglia, Norwich). We're interested in this because there is evidence to suggest that food can affect our gut bacteria in a positive way. We know that orange juice provides a good source of gut friendly chemicals (called polyphenols), hence we would like to examine how the different types of orange-based study drinks affect the amount and type of children's gut bacteria over the 10 weeks of consumption.

Eligible children will then receive one of the orange drinks daily for 10 weeks. Each child will be randomly assigned to consuming either drink A or drink B daily. Allocation to each drink will be decided by chance – rather like tossing a coin. There is an equal chance that your child will receive drink A or drink B. University staff will be responsible for delivering the drinks to school to give to the children and for supervising drink consumption during school hours. Drinks will be sent home with each child on weekends and short school holidays in a cool bag for consumption when not at school. We would be very grateful if you could sign a very short drink diary confirming the days your child has had the drink.

During the 10 weeks, children will complete a 30 minute block of cognitive tests on 3 separate days over that period during the school day on the school premises. The cognitive tests are administered using touch screen tablets and most children find them fun. The tasks will include tests of different abilities such as memory and rule learning. These tasks are not intended to be stressful and are appropriate for your child's age group. Your child will have to option to stop the tests and leave the study at any time.

At the end of the 10 weeks, the children will complete the two well-being and mood questionnaires a second time. We would also ask you and your child's teacher to complete the behaviour and well-being questionnaires, for a second time. This is to examine the impact of the orange-based drinks on their behaviour and well-being.

For the gut bacteria bit, we would again ask that your child provides a further small faecal (stool) sample at home following the 10 week orange juice intervention.

An honorarium of £50 in Love2Shop vouchers will be given upon completion of the study.

Further questions:

What will happen to the stool samples my child provides?

The stool samples that your child provides will be frozen and securely stored for subsequent analysis at the Norwich Research Park (NRP) Biorepository which is adjacent to the Norfolk and Norwich University Hospital (NNUH). We will then examine the impact of the orange-based drink that your child consumes on the amount and type of their gut bacteria, as recent evidence suggests that gut bacteria may be important in overall health, including brain function. Any unused samples will be destroyed. Your child's stool samples will only be labelled with their unique study identity code, date and time so they will not be identifiable from these samples. The researchers analysing your child's stool samples will not have access to the link between their name and unique study identity code.

What do the questionnaire scores mean?

It is important to remember that it is expected that your child will be in the low to normal range for the statements on the questionnaires. However, if this does raise concerns for you please contact me (see below).

What support will I get?

If you have any questions or concerns about the study, any problems providing the faecal samples, or any problems completing the questionnaires please contact me (see below). If necessary I can arrange a home visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

Further Details:

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 16-0187 approved 16/07/2016). All of the information collected from your

child during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If you decide to allow your child to take part they are free to withdraw at any time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if your child participates in the study, you can still request the withdrawal of their information up until completion of work for publication (July 2017). If you wish to remove your child's data from the study at a later date please contact one of the research team (details below) or let your child's teacher know. At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. Some results from the study will be used towards an educational qualification by members of the research team and published in international scientific journals.

This study will be ongoing in school during the period December 2016 - April 2017. All researchers will provide full DBS clearance prior to undertaking the study.

Opt-in:

If you are you happy for your child to participate in this study then please inform the school using the slip below. If you do not wish for your child to participate you do not need to do anything further.

Please note: To ensure that children who have an allergy or intolerances to fruit drinks do not take part in this study can you please complete the attached questionnaire if relevant for your child. Information on allergies will also be provided by the school.

If you have any questions about the study please contact:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Dr James Stone: 0113 343 1403 / email: j.m.stone@leeds.ac.uk

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

PLEASE COMPLETE AND RETURN IF YOU DO WISH YOUR CHILD TO TAKE PART

An intervention study of the effects of 100% fresh orange juice on

(1) cognitive performance
(2) behaviour and well-being
(3) gut bacteria
in 8-10 year old school children

I **DO** wish my son/daughter

.....

(Please insert child's name)

to take part in the above research study.

Signed.....(parent/carer)

Date.....

Mobile number :……..(parent/carer).

Please note: We will only use this number if we need to contact you about the study (e.g. to send you text reminders about the questionnaires and/or stool samples).

Please only return this slip if you DO wish your child to take part in this study.

Please reply by 9th December 2016

PLEASE COMPLETE AND RETURN IF YOU DO WISH YOUR CHILD TO TAKE PART

Name of child……..

1. Does your child take any **medication** or **vitamin tablets**?

☐ Yes

☐ No

If yes, please state _____

2. Does your child have any **food allergies** or **food intolerances**? For example, is your child allergic to nuts or fruit or intolerant to milk?

☐ Yes

☐ No

If yes, please state _____

3. Are there any foods your child cannot eat? For example, food that is not **Halal** or **Kosher**.

☐ Yes

☐ No

If yes, please state _____

4. Is your child **unwell** now or have they been ill during the past few days?

☐ Yes

☐ No

If yes, please state what is/was the illness_____

_____

Please reply by……

Now please return your questionnaire back to School

Appendix 23: Thesis Intervention Study, Cohort 2 - Combined Teacher Information Sheet (all study elements combined)

UNIVERSITY OF LEEDS

School of Psychology
Leeds LS2 9JT

Dear Teachers and Teaching Assistants,

An intervention study of the effects of 100% fresh orange juice on
(1) cognitive performance
(2) behaviour and well-being
and
(3) gut bacteria
in 8-10 year old school children

Professor Louise Dye from the University of Leeds is currently leading a team who are conducting a study in your pupil's school. This work is a research collaboration between the University of Leeds and the State of Florida Department of Citrus.

We would like to test children in your class whose parents have consented for them to take part in a study examining how different types of orange-based drinks affect children's cognitive performance and well-being over a period of 10 weeks. This letter is designed to provide you with information about the study.

What the study entails:

At an initial screening, children will be taken out of lessons for 30 minutes to complete a short IQ and colour blindness test. If any children are found to be colour-blind we will pass this information onto the school. They will also be asked to complete a peak flow test to measure lung capacity.

During the screening, the child will also complete two questionnaires asking about their well-being and mood. The two questionnaires are called (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale. These will be completed in a self-report format by the children, with assistance from the research staff. The first one is about assessing the children's emotional and behavioural strengths and

difficulties. The second is about assessing signs of anxiety and depression in the children. We would like to examine how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks as there is some evidence to suggest that healthy foods can affect the well-being and behaviour of children in a positive way.

Following the initial screening, we would ask you (as their form teacher) and your pupils' parent/guardians to **complete a total of three questionnaires** to examine your child's behaviour and well-being prior to receiving one of the orange drinks. These will include (1) The Strengths and Difficulties Questionnaire and (2) The Revised Child Anxiety and Depression Scale, in addition to (3) the SNAP-IV. The additional questionnaire is about assessing the children's' attention and activity levels. As their form teacher you are ideally placed to give us day-to-day knowledge about your pupil. **The questionnaires require you to read a series of statements and judge how well it describes the child by ticking one of the options. Completion shouldn't take long – about 10 minutes on average**

For the gut bacteria bit, we are also asking that that the child provides a small faecal (stool) sample at home before they start receiving one of the orange-based drinks. We will provide the parent/guardians with a specially designed stool collection kit, instructions on how to use this kit and freepost postal boxes (one for each sample) so that they can post the samples off for analysis. This analysis will be carried out by our collaborator (Dr David Vauzour from the University of East Anglia, Norwich). We're interested in this because there is evidence to suggest that food can affect our gut bacteria in a positive way. We know that orange juice provides a good source of gut friendly chemicals (called polyphenols) hence we would like to examine how the different types of orange-based study drinks affect the amount and type of children's gut bacteria over the 10 weeks of consumption.

Eligible children will then receive one of the orange drinks daily for 10 weeks. Each child will be randomly assigned to consuming either drink A or drink B daily. Allocation to each drink will be decided by chance – rather like tossing a coin. There is an equal chance that the children will receive drink A or drink B. University staff will be responsible for delivering the drinks to school to give to the children and for supervising drink consumption during school hours. Drinks will be sent home with each child on weekends and short school holidays in a cool bag for consumption when not at school. The parent/guardians should also sign a very short drink diary confirming the days the child has had the drink.

During the 10 weeks, children will be taken out of class to complete a 30 minute block of cognitive tests on 3 separate days over that period, on the school premises. The cognitive tests are administered using touch screen tablets and most children find them fun. The tasks will include tests of different abilities such as memory and rule learning. These tasks are not

intended to be stressful and are appropriate for the children's age group. The children will have the option to stop the tests and leave the study at any time.

At the end of the 10 weeks, the children will complete the two well-being and mood questionnaires a second time. We would also ask you and your child's parent/guardians to **complete the behaviour and well-being questionnaires, for a second time**. This is to examine the impact of the orange-based drinks on their behaviour and well-being.

For the gut bacteria bit, we are again asking that the child provides a further small faecal (stool) sample at home following the 10 week orange juice intervention.

An honorarium of £50 in Love2Shop vouchers will be given to the families upon completion of the study.

Further questions:

What will happen to the stool samples the children provide?

The stool samples that the children provide will be frozen and securely stored for subsequent analysis at the Norwich Research Park (NRP) Biorepository which is adjacent to the Norfolk and Norwich University Hospital (NNUH). We will then examine the impact of the orange-based drink that the children consumes on the amount and type of their gut bacteria, as recent evidence suggests that gut bacteria may be important in overall health, including brain function. Any unused samples will be destroyed. The children's stool samples will only be labelled with their unique study identity code, date and time so they will not be identifiable from these samples. The researchers analysing the stool samples will not have access to the link between their name and unique study identity code.

What do the questionnaire scores mean?

It is important to remember that it is expected that the children will be in the low to normal range for the statements on the questionnaires. However, if this does raise concerns for you please contact me (see below).

What support will I get?

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me (see below). If necessary I can arrange a school visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

Further Details:

This study has received ethical approval from the School of Psychology Research Ethics Committee (Ref 16-0187 approved 16/07/2016). All of the information collected from the children during the study will be kept strictly confidential and will only be used for the purposes of this research. Participation in the study is completely voluntary. If parent/guardians decide to allow their children to take part they are free to withdraw at any

time without providing a reason. All results from the study will be kept strictly anonymous and at no point will any identifiable personal information be linked with the results. Even if a child participates in the study, they can still request the withdrawal of their information up until completion of work for publication (July 2017). If you wish to remove a child's data from the study at a later date please contact one of the research team (details below). At the end of the study both teachers and parents will be invited to a post study dissemination event where top line results will be presented by us and our supervisors. Some results from the study will be used towards an educational qualification by members of the research team and published in international scientific journals.

This study will be ongoing in school during the period January - April 2017. All researchers will provide full DBS clearance prior to undertaking the study.

If you have any questions about the study please contact:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Dr James Stone: 0113 343 1403 / email: j.m.stone@leeds.ac.uk

Alternatively you can email or phone the project supervisors Professor Louise Dye and Dr Clare Lawton at the School of Psychology:

Prof Louise Dye: 0113 3435707 or l.dye@leeds.ac.uk

Dr Clare Lawton: 0113 3435741 or c.l.lawton@leeds.ac.uk

Appendix 24: Thesis Intervention Study, Cohort 2 - Combined Child Information/Consent Sheet (All study elements combined)

UNIVERSITY OF LEEDS

Hi there!

We are researchers from the University of Leeds.

We want to know if what you eat and drink has an effect on your memory and attention, or maybe even your mood and behaviour.

We are also interest to see how what you eat and drink affects the bacteria in your gut.

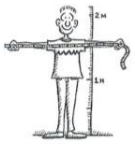
To find this out, we would like you to take part in our study!

What you will need to do:

- Have orange juice every morning for 10 weeks
- Do some computer tests on touchscreen tablets
- Answer some questions about yourself
- Provide a stool sample at the start and end of the study.

In <u>week 1</u>, we will find out whether you can join our study by doing:

- A colour blindness test



- Two short puzzles

- Measure your height and weight

- Have you blow into a tube 3 times to measure how strong your lungs are

- Answer some questionnaires about your mood

- Send off a stool sample

<u>In week 2</u>, you will

- Start drinking orange juice everyday for 10 weeks. We will give you bottles to drink at home at the weekend too

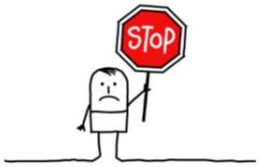
- Complete more computer-based tasks after you have had your first orange juice drink

<u>After 10 weeks</u>

- You will repeat the computer-based tasks

- You will repeat the questionnaires about your mood

You do not have to take part if you don't want to. That is okay. Just tell us or your teacher



If you do take part, you can change your mind at any time and stop without telling us why. Just tell us or your teacher if you want to stop.

We will use your results to write reports for other researchers. Nobody will know your name, only a number that is linked to your results so no one can tell the information came from you.

If you have any questions or if you are unsure about anything, please ask us.

You do NOT have to decide RIGHT NOW if you want to take part. Take you time to think about it.

If you do not want to take part, that is okay. You can go back to class

If you do want to take part we can start

If you understand what the study is about and are happy to continue, please sign your name below

_____

Appendix 25: Thesis Intervention Study, Cohort 2 - Short information/reminder sheet sent to parents prior to onset of intervention.

School of Psychology
Leeds LS2 9JT

Dear Parent/Caregiver,

We recently sent you information about a study we are running in your child's school which is looking at the effects of daily orange juice consumption on cognitive performance, behaviour and wellbeing and gut bacteria. You are receiving this letter because you agreed to let your child participate in this. We would like to thank you very much for this.

Your child recently began the study and will soon be starting to consume one bottle of provided orange juice from Florida per day.

As part of the study we now need your help. As mentioned in the information sheet you saw previously we would like you to complete some short questionnaires about your child and send a stool sample from your child **before your child starts drinking the study orange juice daily**.

We would really appreciate it if you could complete the following activities before xx/xx/xxxx.

1. We now ask you to fill in the three questionnaires included with this letter to see your views about your child's behaviour and well-being before they get any orange juice. These include (1) The Strengths and Difficulties Questionnaire, (2) The Revised Child Anxiety and Depression Scale, and (3) the SNAP-IV. As their parent/guardian you are the best person to give us day-to-day knowledge about your child. All the questionnaires ask you to read a series of sentences and judge how well each describes your child by ticking one of the options. It shouldn't take long – about 15 minutes.
2. We have included a specially designed stool collection kit with this letter. There are instructions on how to use the kit, to collect the sample. There are also

instructions on how to use the freepost postal boxes (one for each sample), so that you can post your child's samples off for analysis.

At the end of the study: After 10 weeks of daily orange juice consumption **we will also ask you to complete the behaviour and well-being questionnaires and send a stool sample, for a second time**. This is to see if the drinks made any difference to your child's behaviour and wellbeing.

Reminder: A reward of £50 in Love2Shop vouchers will be given upon completion of the full study.

Contact details:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Appendix 26: Thesis Intervention Study, Cohort 2 - Short information/reminder sheet sent to parents at the end of intervention

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

Dear Parent/Caregiver,

Your child has now completed our orange juice study!

As a final step we now need your help again. As you have done previously (at the start of the study) we would like you to complete some short questionnaires about your child and send a stool sample from your child **as soon as possible**.

We would really appreciate it if you could complete the following activities before xx/xx/xxxx.

1. We now ask you to fill in the three questionnaires included with this letter to see your views about your child's behaviour and well-being. These include (1) The Strengths and Difficulties Questionnaire, (2) The Revised Child Anxiety and Depression Scale, and (3) the SNAP-IV. As their parent/guardian you are the best person to give us day-to-day knowledge about your child. All the questionnaires ask you to read a series of sentences and judge how well each describes your child by ticking one of the options. It shouldn't take long – about 15 minutes.
2. We have included a specially designed stool collection kit with this letter. There are instructions on how to use the kit, to collect the sample. And there are also instructions on how to use the freepost postal boxes (one for each sample), so that you can post your child's samples off for analysis.

Reminder: A reward of £50 in Love2Shop vouchers will be given upon completion of the full study.One last time we would like to thank you very much for your cooperation throughout this study.

Contact details:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Appendix 27: Thesis Intervention Study, Cohort 2 - Short information/reminder sheet sent to teachers prior to onset of intervention.

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

An intervention study of the effects of 100% fresh orange juice on cognitive performance, behaviour and well-being and gut bacteria in 8-10 year old children

My name is Fiona Wright, I am a Psychologist in Clinical Training from the University of Leeds. I am part of the team who are currently conducting a study in your pupil's school. As part of this study, we are examining how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks. There is some evidence to suggest that healthy foods can affect the well-being and behaviour of children in a positive way.

As part of the study, we are asking you to complete two questionnaires at two time points (10 weeks apart), for each of the children in your class who are taking part in the study. This letter is designed to provide you with information about completing the first set of behaviour and well-being questionnaires. It is important that you complete these questionnaires **before the children start drinking the study orange juice**. So we ask that you return the questionnaires within the next two weeks (by xx/xx/xxxx at the latest).

For each participating child in your class, we now ask that you complete the two questionnaires included with this letter to assess your perspective about your pupil's behaviour and well-being **before** they receive any orange juice. These are (1) The **Strengths and Difficulties Questionnaire** and (2) the **SNAP-IV**. The first one is about assessing your pupil's emotional and behavioural strengths and difficulties. The second is about assessing your pupil's attention and activity levels.

As their teacher you are ideally placed to give us day-to-day knowledge about your pupils. All the questionnaires ask you to read a series of statements and judge how well it describes your pupil by ticking one of the options. It shouldn't take long – no more than 10 minutes.

At the end of the study:

At the end of the 10 week study, your pupil will complete the two well-being and mood questionnaires a second time. At this time point, **we will also ask you to complete the behaviour and well-being questionnaires on each participating child, for a second time**. This is to examine the impact of the orange-based drinks on their behaviour and well-being.

Support:

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me (see below). If necessary I can arrange a school visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that your pupils will all be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please contact me (see below).

An honorarium of £2 in Love2Shop vouchers will be payable to you for each full set of study questionnaires that you return for each child (i.e. 2 questionnaires at the start and end of the study per child).

Contact details:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Appendix 28: Thesis Intervention Study, Cohort 2 - Short information/reminder sheet sent to teachers at the end of the intervention.

UNIVERSITY OF LEEDS

School of Psychology

Leeds LS2 9JT

An intervention study of the effects of 100% fresh orange juice on cognitive performance, behaviour and well-being and gut bacteria in 8-10 year old children

My name is Fiona Wright, I am a Psychologist in Clinical Training from the University of Leeds. I am part of the team who are currently conducting a study in your pupil's school. As part of this study, we are examining how different types of orange-based drinks affect children's behaviour and well-being over a period of 10 weeks. There is some evidence to suggest that healthy foods can affect the well-being and behaviour of children in a positive way.

As part of the study, we are asking you to complete two questionnaires at two time points (10 weeks apart), for each of the children in your class who are taking part in the study. This letter is designed to provide you with information about completing the final set of behaviour and well-being questionnaires. It is important that you complete these questionnaires **as soon as possible after the children have finished the 10-weeks of daily orange juice consumption**. So we ask that you return the questionnaires within the next two weeks (by xx/xx/xxxx at the latest).

For each participating child in your class, we now ask that you complete the two questionnaires included with this letter to assess your perspective about your pupil's behaviour and well-being **before** they receive any orange juice. These are (1) The **Strengths and Difficulties Questionnaire** and (2) the **SNAP-IV**. The first one is about assessing your pupil's emotional and behavioural strengths and difficulties. The second is about assessing your pupil's attention and activity levels.

As their teacher you are ideally placed to give us day-to-day knowledge about your pupils. All the questionnaires ask you to read a series of statements and judge how well it describes your pupil by ticking one of the options. It shouldn't take long – no more than 10 minutes.

Support:

If you have any questions or concerns about the study or any problems completing the questionnaires please contact me (see below). If necessary I can arrange a school visit to discuss the study with you and answer any questions. Alternatively we can talk it through over the phone.

What do the questionnaire scores mean?

It is important to remember that it is expected that your pupils will all be in the low to normal range for the statements on the questionnaires. However, if this raises any concerns for you then please contact me (see below).

An honorarium of £2 in Love2Shop vouchers will be payable to you for each full set of study questionnaires that you return for each child (i.e. 2 questionnaires at the start and end of the study per child).

Contact details:

Dr Fiona Wright: 07956024807 / email: ps07fw@leeds.ac.uk

Appendix 29: SDQ Teacher Informant Data - Factor Structure

	Three Factor Solution			Four Factor Solution				Five Factor Solution				
	1	2	3	1	2	3	4	1	2	3	4	5
21. REFLECTIVE(R)_H	0.74			.727						-0.70		
25. PERSISTENT(R)_H	0.65			.716						-0.70		
15. DISTRACTABLE_H	0.67		0.38	.703		.348				-0.70		
7. OBEDIENT(R)_C	0.52			.449						-0.42		
5. TEMPERS_C	0.40			.342						-0.33		
12. FIGHTS_C		0.41			.854				0.84			
22. STEALS_C		0.51			.721				0.73			
18. LIES_C		0.34			.569				0.60			
9. CARING_PS		-0.39			-.384				0.37			
1. CONSIDERATE_PS	-0.42			-.308	-.382				-0.41	0.33		
20. HELPFUL_PS	-0.33			-	-	-	-	-	-	-	-	-
10. FIDGETY_H			0.79			.798		0.95				
2. RESTLESS_H			0.62	.448		.612		0.64				
8. WORRIES_E			0.54			.537						0.44
24. FEARS_E			0.46			.459						0.58
16. CLINGY			0.42			.423		-	-	-	-	-
3. SOMATIC_E	-	-	-	-	-	-	-					0.37
11. GOOD FRIEND(R)_PP		0.43					.611				0.63	
14. POPULAR(R)_PP		0.52					.583				0.78	
23. BETTER WITH ADULTS_PP		0.64					.547				0.45	
19. BULLIED_PP		0.56					.537				0.45	
6. SOLITARY_PP		0.37					.502				0.42	0.31
13. UNHAPPY_E		0.50					.402					0.39
17. KIND_PS	-0.33	-0.32					-.312			0.33		
4. SHARES_PS	-	-	-	-	-	-	-	-	-	-	-	-

Table *:** ***Factor loadings and communalities based on a maximum likelihood factor analysis with oblimin rotation for 25 items from the strengths and difficulties questionnaire (SDQ; N = 101).****Note: Correlations below 0.3 are not included. Red text indicates items that are loading on multiple factors. Dashes (-) indicate variables that fail to load onto any factor*

Appendix 30: The Means of the SDQ Total Difficulty Composite Score and Sub-scale Scores in Different Studies

Studies	Population	Samples	Age range	Gender	Total Difficulty	Emotion	Conduct	Hyperactivity	Peer problems	Prosocial
Teacher-Informed										
Koskelainen et al. (2000) *	Finish	n = 376	7 to 12	Boys	7.8					
				Girls	4.8					
Meltzer et al. (2000)	British	n = 4,801	5 to 10	Boys	8.0					
				Girls	5.6					
		n = 3,407	11 to 15	Boys	7.6					
				Girls	5.0					
Capron et al. (2007)	French	n = 1,400	11 to 16	Boys	9.5	1.6	1.6	4.2	2.1	6.1
				Girls	6.9	2.0	0.7	2.5	1.7	7.1
d'Acremont and Linden (2008)	Swiss	n = 557	13 to 18	Total		1.8	1.1	2.6	1.8	1.8
Du et al. (2008)	Chinese	n = 1,217	3 to 10	Boys	10.6	1.8	1.7	4.7	2.5	6.3
				Girls	8.0	1.8	1.0	3.2	2.0	7.3
		n = 748	11 to 17	Boys	11.0	1.8	1.7	4.5	2.9	6.2
				Girls	8.3	1.9	1.2	2.8	2.4	7.7
D. Mellor (2004)	Australian	n = 917	7 to 17	Total	6.5	1.4	1.0	2.5	1.6	7.8
Parent-Informed										
Koskelainen et al. (2000) *	Finish	n = 703	7 to 12	Boys	6.7					
				Girls	6.2					
			13 to 16	Boys	5.3					
				Girls	5.3					
Meltzer et al. (2000)	British	n = 5,055	5 to 10	Boys	9.3					
				Girls	7.9					

		n = 4,443	11 to 15	Boys	8.8					
				Girls	7.6					
Woerner et al. (2004)	German	n = 930	6 to 16	Boys	8.53	1.5	1.95	3.40	1.67	7.55
				Girls	7.72	1.57	1.68	2.97	1.51	7.67
Du et al. (2008)	Chinese	n = 1,217	3 to 10	Boys	11.3	1.9	1.8	4.9	2.7	6.8
				Girls	10.5	2.3	1.4	4.1	2.7	7.5
		n = 748	11 to 17	Boys	11.0	1.7	1.7	4.3	3.1	6.8
				Girls	8.3	1.9	1.4	3.3	2.4	7.5
Mellor (2004)	Australian	n = 917	7 to 17	Total	8.2	2.1	1.5	3.1	1.6	8.3
Bull et al. (2016)	Singaporean	n = 411	3 to 16	Boys	10.7	1.7	2.5	4.5	2.1	6.4
				Girls	8.9	1.8	1.9	3.4	2.0	6.8
Lai et al. (2010)	Hong Kong		6 to 12	Boys	12.3	2.3	2.3	4.9	2.7	6.4
				Girls	11.2	2.6	2.0	4.2	2.5	7.0
Mieloo et al. (2012)	Netherlands		5 to 6	Boys	6.7	1.4	1.3	3.1	0.9	8.0
				Girls	5.5	1.4	1.0	2.3	0.7	8.6
Mansbach-Kleinfeld et al. (2010)	Israeli / Hebrew	n = 533	14 to 17	Boys	6.7	1.6	1.5	2.2	1.5	8.5
				Girls	3.4	2.0	1.4	1.8	1.3	8.8
Self-Reported										
Koskelainen et al. (2000) *	Finish	n = 528	7 to 12	Boys	9.6					
				Girls	8.3					
			13 to 16	Boys	7.5					
				Girls	7.0					
Koskelainen et al. (2001)	Finish	n = 1,458	13 to 14	Boys	10.5	2.2	2.5	3.4	2.5	6.1
				Girls	11.6	3.4	2.2	3.6	2.4	7.1
			15 to 17	Boys	11.0	2.4	2.6	3.5	2.5	5.9
				Girls	11.8	3.6	2.3	3.8	2.2	7.5

Ronning, Handegaard, Sourander, and Morch (2004)	Norwegian	n = 2,028	13 to 14	Boys	9.6	2.0	2.0	3.7	2.0	6.8
				Girls	8.8	2.6	1.5	3.1	1.7	7.9
			15 to 17	Boys	9.9	1.9	2.3	4	1.8	6.1
				Girls	10.1	3.1	1.6	3.8	1.5	7.8
Van Roy, Groholt, Heyerdahl, and Clench-Aas (2006)	Norwegian	n = 9,707	10 to 13	Boys	10.3	2.2	2.1	3.8	2.1	7.4
				Girls	10.1	3.0	1.7	3.5	1.9	8.2
		n = 9,387	13 to 14	Boys	10.8	2.1	2.5	4.2	2.1	6.5
				Girls	11.1	3.2	2.0	4.2	1.8	7.7
Meltzer et al. (2000)	British	n = 4,228	11 to 15	Boys	10.5	2.6	2.4	3.9	1.6	7.5
				Girls	10.0	3.0	2.0	3.6	1.4	8.5
Capron et al. (2007)	French	n = 1,400	11 to 16	Boys	11.0	2.4	2.8	3.8	2.0	7.4
				Girls	11.5	3.7	2.4	3.5	1.9	8.2
Di Riso et al. (2010)	Italian	n = 1,394	8 to 10	Boys	15.0	3.5	3.6	4.6	3.3	6.4
				Girls	15.0	4.4	3.2	4.0	3.3	7.0
Giannakopoulos et al. (2009)	Greek	n = 1,194	11 to 14	Boys		2.3	3.0	3.1	1.9	7.9
				Girls		3.1	2.8	3.1	1.7	8.4
			15 to 17	Boys		2.5	3.0	3.9	1.9	7.5
				Girls		3.7	3.1	4.1	2.0	8.2
Lundh et al. (2008)	Swedish	n = 504	Grade 7	Boys	9.9	2.0	1.9	3.9	2.2	6.9
				Girls	10.4	3.3	1.6	3.7	1.8	8.2
		n = 482	Grade 8	Boys	9.8	1.8	2.2	3.8	1.9	6.8
				Girls	10.4	3.2	1.8	3.8	1.7	8.1
Du et al. (2008)	Chinese	n = 960	11 to 17	Boys	11.5	2.4	2.4	3.7	3.2	6.8
				Girls	9.8	2.3	2.0	3.0	2.5	7.8
Mellor and Stokes (2007)	Australian	n = 359	7 to 11	Total	9.7	3.1	1.7	3.0	1.9	8.4
		n = 558	11 to 17	Total	9.0	2.4	1.8	3.2	1.5	8.0

Mansbach-Kleinfeld et al. (2010)	Israeli / Hebrew	n = 533	14 to 17	Boys	8.2	2.1	1.8	2.9	1.5	8.0
				Girls	9.3	1.9	2.0	3.0	2.3	8.2
Yao et al. (2009)	Chinese	n = 394	11 to 14	Boys	10.6	1.9	2.4	3.3	3.0	5.0
				Girls	9.4	2.0	1.9	3.1	2.4	5.1
		n = 741	15 to 18	Boys	10.1	1.8	2.2	3.7	1.7	2.7
				Girls	10.6	2.4	2.1	3.8	2.3	6.6

Appendix 31: The Inter-rater Agreement (Pearson's product moment correlation co-efficient) of the SDQ scales in different studies

Studies	Population	Sample size	Age range	Total Difficulty	Emotion	Conduct	Hyperactivity	Peer problems	Prosocial
Parent/Teacher									
Koskelainen et al. (2000)	Finnish	n = 376	7 to 12	0.44	0.33	0.3	0.45	0.39	0.29
Goodman (1997)	English	n = 128	4 to 16	0.62	0.41	0.65	0.54	0.59	0.37
Goodman et al. (1998)	English	n = 83	11 to 16	0.43	0.38	0.28	0.36	0.14	0.16
Goodman (2001)	English	n = 7,313	5 to 15	0.46	0.27	0.37	0.48	0.37	0.25
Du et al. (2008)	Chinese	n = 1,965	3 to 17	0.36	0.23	0.31	0.44	0.29	0.27
Mellor (2004)	Australian	n = 917	7 to 17	0.46	0.31	0.34	0.46	0.39	0.30
Parent /Self-report									
Koskelainen et al. (2000)	Finnish	n = 528	9 to 16	0.40	0.28	0.28	0.39	0.39	0.37
Goodman et al. (1998)	English	n = 83	11 to 16	0.43	0.52	0.36	0.29	0.29	0.31
Goodman (2001)	English	n = 3,983	11 to 15	0.48	0.37	0.44	0.41	0.40	0.30
Du et al. (2008)	Chinese	n = 690	3 to 17	0.49	0.72	0.39	0.37	0.36	0.40
Mellor (2004)	Australian	n = 917	7 to 17	0.45	0.32	0.37	0.46	0.34	0.45
Teacher/Self-report									
Koskelainen et al. (2000)	Finnish	n = 376	9 to 16	0.38	0.25	0.3	0.34	0.38	0.28
Goodman et al. (1998)	English	n = 83	11 to 16	0.38	0.31	0.19	0.13	0.24	0.19
Goodman (2001)	English	n = 2,767	11 to 16	0.33	0.21	0.3	0.32	0.29	0.23
Du et al. (2008)	Chinese	n = 690	3 to 17	0.42	0.29	0.34	0.38	0.35	0.31
Mellor (2004)	Australian	n = 917	7 to 17	0.45	0.24	0.39	0.44	0.35	0.29

Appendix 32: The Internal Consistency (Cronbach's Alpha) of the SDQ Scales in Different Studies

Studies	Population	Sample size	Age range	Total Difficulty	Emotion	Conduct	Hyperactivity	Peer problems	Prosocial
Teacher Informed									
Koskelainen et al. (2000)	Finnish	n = 376	7 to 12	0.71	0.79	0.72	0.85	0.73	0.86
Goodman (2001)	British	n = 7,313	5 to 15	0.8	0.65	0.69	0.82	0.72	0.74
Capron et al. (2007)	French	n = 1,400	11 to 16	0.83	0.78	0.76	0.87	0.64	0.83
d'Acremont and Linden (2008)	Swiss	n = 557	13 to 18		0.27	0.14	0.00	0.19	-0.07
Du et al. (2008)	Chinese	n = 1,965	3 to 17	0.60	0.63	0.63	0.82	0.48	0.83
Mellor (2004)	Australian	n = 917	7 to 17	0.76	0.77	0.75	0.87	0.71	0.83
Parent-Informed									
Koskelainen et al. (2000)	Finnish	n = 703	7 to 16	0.71	0.69	0.59	0.73	0.64	0.68
Smedje et al. (1999)	Swedish	n = 900	6 to 10	0.76	0.61	0.54	0.75	0.51	0.7
Goodman (2001)	British	n = 9,998	5 to 15	0.82	0.67	0.63	0.77	0.57	0.65
Malmberg et al. (2003)	Swedish	n = 263	5 to 15	0.84	0.71	0.52	0.75	0.73	0.67
Muris et al. (2003)	Dutch	n = 562	9 to 15	0.8	0.7	0.55	0.78	0.66	0.68
Woerner et al. (2004)	German	n = 930	6 to 16	0.82	0.66	0.6	0.76	0.58	0.68
Du et al. (2008)	Chinese	n = 1,965	3 to 17	0.59	0.60	0.48	0.76	0.30	0.68
Mellor (2004)	Australian	n = 917	7 to 17	0.73	0.71	0.67	0.80	0.75	0.70
Self-Reported									
Koskelainen et al. (2000)	Finnish	n = 528	9 to 16	0.71	0.69	0.57	0.66	0.63	0.69
Goodman et al. (1998)	English	n = 83	11 to 16	0.82	0.75	0.72	0.69	0.61	0.65
Goodman (2001)	English	n = 3,983	11 to 15	0.8	0.66	0.6	0.67	0.41	0.66
Muris et al. (2003)	Dutch	n = 562	9 to 15	0.78	0.71	0.45	0.72	0.54	0.62
Capron et al. (2007)	French	n = 1,400	11 to 16	0.73	0.64	0.53	0.66	0.46	0.59
Di Riso et al. (2010)	Italian	n = 1,394	8 to 10	0.70	0.55	0.41	0.57	0.27	0.67

Lundh et al. (2008)	Swedish	n = 926	14 to 15	0.76	0.67	0.57	0.66	0.56	0.68
Du et al. (2008)	Chinese	n = 960	11 to 17	0.57	0.59	0.33	0.64	0.30	0.66
Mellor (2004)	Australian	n = 917	7 to 17	0.71	0.67	0.65	0.74	0.58	0.64
Yao et al. (2009)	Chinese	n = 1,135	11 to 18	0.81	0.78	0.67	0.73	0.52	0.87

Appendix 33: The Factor Structure of the SDQ in Different Studies

Authors	Analysis	Population	Sample Size	Age range	Informants	SDQ Factor Structure
Teacher-Reported						
d'Acremont and Linden (2008)	CFA	Swiss	n = 557	13 to 18	Teachers	5 factor (?)
Du et al. (2008)	CFA	Chinese Community	n = 1,965	3 to 17	Teachers	5 Factor (?)
Mellor and Stokes (2007)	CFA	Australian	n = 914	7 to 17	Teachers	5 Factor (-)
Parent-Reported						
Capron et al. (2007)	FA	French	N = 1,400	11 to 16	Parents	5 Factor
Smedje et al. (1999)	PCA	Swedish	n = 900	6 to 10	Parents	5 factor
W. Woerner et al. (2002)		German	n = 930	6 to 16	Parents	5 factor
Mansbach-Kleinfeld et al. (2010)	EFA / CFA	Israeli / Hebrew	n = 533	14 to 17	Parents	5 Factor (?)
Du et al. (2008)	CFA	Chinese Community	n = 1,965	3 to 17	Parents	5 Factor (?)
Mellor, Wong, and Xu (2011)	CFA	Chinese Clinical			Parents	5 Factor (?)
Mellor and Stokes (2007)	CFA	Australian	n = 914	7 to 17	Parents	5 Factor (-)
Bull et al. (2016)	CFA	Singaporean	n = 411	3 to 16	Parents	4 Factor
Dickey and Blumberg (2004)	PCA	US	n = 9,574	4 to 17	Parents	3 factor
Self-Reported						
Capron et al. (2007)	PCA	French	N = 1,400	11 to 16	Self-Report	5 factor
R. Goodman (2001)	PCA	English	n = 3,983	11 to 15	Self-Report	5 factor
Muris et al. (2004)	PCA	Dutch	n = 562	9 to 15	Self-Report	5 factor
Koskelainen et al. (2000)	PCA	Finnish	n = 528	9 to 16	Self-Report	5 factor
Van Roy et al. (2008)	CFA	Norwegian	n = 26,269	10 to 19	Self-Report	5 factor
Giannakopoulos et al. (2009)	CFA	Greek	n = 1,194	11 to 17	Self-Report	5 Factor
Lundh et al. (2008)	PCA	Swedish	n = 1254	14 to 15	Self-Report	5 factor

Yao et al. (2009)	CFA	Chinese	n = 1,135	11 to 18	Self-Report	5 Factor
Ronning et al. (2004)	CFA	Norwegian	n = 4,167	11 to 16	Self-Report	5 Factor (?)
Mansbach-Kleinfeld et al. (2010)	EFA / CFA	Israeli / Hebrew	n = 611	14 to 17	Self-Report	5 Factor (?)
Du et al. (2008)	CFA	Chinese Community	n = 1,965	3 to 17	Self-Report	5 Factor (?)
Mellor and Stokes (2007)	CFA	Australian	n = 914	7 to 17	Self-Report	5 Factor (-)
Koskelainen et al. (2001)	PFA	Finish	n = 1,458	13 to 17	Self-Report	3 factor
Percy et al. (2008)	CFA	Northern Ireland	n = 3,753	Age 12	Self-Report	3 factor
Ruchkin et al. (2008)	CFA	American Urban	n = 4,661	11 to 16	Self-Report	3 factor
Di Riso et al. (2010)	CFA	Italian	n = 1,394	8 to10	Self-Report	3 Factor
Palmieri and Smith (2007)	CFA	US custodial	n = 733	4 to 16	Grandparents	4 factor
Mellor et al. (2011)	CFA	Chinese Clinical			Self-Report	
Haynes et al. (2013)	EFA / CFA	Australian Intellectual Disability	n = 128	9 to 14	Self-Report	3 Factor

Appendix 34: The Main Effects and Interactions for SDQ Total Difficulty and Prosocial Strength Sub-scales

	SDQ Total Difficulty			SDQ Prosocial Strengths		
Interaction	***Teacher Response***	***Parents Response***	***Self-Report***	***Teacher Response***	***Parents Response***	***Self-Report***
Treatment x Timepoint	$F(1, 28^{ggadj}) = 0.16, p = 0.69$	$F(1, 27^{ggadj}) = 0.82, p = 0.37$	$F(1, 47^{ggadj}) = 0.26, p = 0.61$	$F(1, 34^{ggadj}) = 0.31, p = 0.58$	$F(1, 24^{ggadj}) = 0.20, p = 0.66$	$F(1, 47^{ggadj}) = 0.11, p = 0.75$
Timepoint x Gender	$F(1, 25^{ggadj}) = 0.61, p = 0.44$	$F(1, 28^{ggadj}) = 0.02, p = 0.86$	$F(1, 47^{ggadj}) = 0.73, p = 0.40$	$F(1, 33^{ggadj}) = 0.40, p = 0.53$	$F(1, 23^{ggadj}) = 0.19, p = 0.67$	$F(1, 47^{ggadj}) = 0.10, p = 0.76$
Timepoint x Age	$F(1, 26^{ggadj}) = 0.12, p = 0.73$	$F(1, 28) = 1.23, p = 0.28$	$F(1, 47^{ggadj}) = 0.01, p = 0.92$	$F(1, 34^{ggadj}) = 0.37, p = 0.55$	$F(1, 22^{ggadj}) = 0.28, p = 0.60$	$F(1, 47^{ggadj}) = 0.05, p = 0.82$
Timepoint x IQ	$F(1, 25^{ggadj}) = 0.001, p = 0.98$	$F(1, 27^{ggadj}) = 0.61, p = 0.44$	$F(1, 47^{ggadj}) = 0.45, p = 0.51$	$F(1, 34^{ggadj}) = 0.01, p = 0.94$	$F(1, 24^{ggadj}) = 1.23, p = 0.28$	$F(1, 47^{ggadj}) = 1.23, p = 0.27$
Treatment x Gender	$F(1, 24^{ggadj}) = 2.32, p = 0.14$	$F(1, 28^{ggadj}) = 0.76, p = 0.39$	$F(1, 47^{ggadj}) = 0.73, p = 0.40$	$F(1, 34^{ggadj}) = 2.85, p = 0.10$	$F(1, 23^{ggadj}) = 0.18, p = 0.67$	$F(1, 47^{ggadj}) = 0.79, p = 0.37$
Treatment x Age	$F(1, 24^{ggadj}) = 0.60$	**$F(1, 28^{ggadj}) = 14.55, p \leq 0.001$**	$F(1, 47^{ggadj}) = 0.07, p = 0.79$	$F(1, 16^{ggadj}) = 0.01, p = 0.94$	$F(1, 22^{ggadj}) = 3.12, p = 0.09$	$F(1, 47^{ggadj}) = 0.48, p = 0.49$
Treatment x IQ	$F(1, 31^{ggadj}) = 2.12, p = 0.15$	$F(1, 26^{ggadj}) = 2.23, p = 0.15$	$F(1, 47^{ggadj}) = 3.50, p = 0.07$	**$F(1, 32^{ggadj}) = 6.88, p \leq 0.01$**	**$F(1, 25^{ggadj}) = 8.76, p < 0.01$**	**$F(1, 47^{ggadj}) = 4.92, p < 0.05$**
Treatment x Timepoint x Gender	$F(1, 24^{ggadj}) = 0.03, p = 0.87$	$F(1, 28^{ggadj}) = 1.95, p = 0.17$	$F(1, 48^{ggadj}) = 0.81, p = 0.37$	$F(1, 34^{ggadj}) = 0.30, p = 0.59$	$F(1, 22^{ggadj}) = 0.62, p = 0.44$	$F(1, 47^{ggadj}) = 2.68, p = 0.11$
Main Effects						
Treatment	$F(1, 24^{ggadj}) = 1.33, p = 0.26$	**$F(1, 27^{ggadj}) = 9.83, p \leq 0.01$**	$F(1, 47^{ggadj}) = 0.12, p = 0.73$	$F(1, 16^{ggadj}) = 1.09, p = 0.31$	**$F(1, 23^{ggadj}) = 8.42, p \leq 0.01$**	$F(1, 47^{ggadj}) = 0.02, p = 0.89$
Timepoint	$F(1, 27^{ggadj}) = 0.18, p = 0.68$	$F(1, 28) = 0.59, p = 0.45$	$F(1, 47^{ggadj}) = 0.14, p = 0.71$	$F(1, 34^{ggadj}) = 0.34, p = 0.56$	$F(1, 22^{ggadj}) = 0.01, p = 0.93$	$F(1, 47^{ggadj}) = 0.05, p = 0.82$
Gender	**$F(1, 25^{ggadj}) = 6.83, p < 0.05$**	**$F(1, 27^{ggadj}) = 9.79, p \leq 0.01$**	**$F(1, 48^{ggadj}) = 9.62, p \leq 0.01$**	**$F(1, 33^{ggadj}) = 7.05, p \leq 0.01$**	$F(1, 25^{ggadj}) = 2.95, p = 0.09$	$F(1, 48^{ggadj}) = 1.41, p = 0.24$
Covariate Effects						
IQ	$F(1, 28.78) = 0.16, p = 0.69$	$F(1, 27.56) = 2.62, p = 0.12$	$F(1, 47.83) = 0.05, p = 0.82$	$F(1, 33^{ggadj}) = 1.77, p = 0.19$	$F(1, 26^{ggadj}) = 0.78, p = 0.39$	$F(1, 48^{ggadj}) = 0.46, p = 0.50$
Age	**$F(1, 25.26) = 11.46, p < 0.01$**	**$F(1, 27.95) = 6.11, p \leq 0.05$**	$F(1, 47.84) = 0.23, p = 0.64$	$F(1, 31^{ggadj}) = 2.83, p = 0.10$	**$F(1, 23^{ggadj}) = 6.27, p \leq 0.05$**	$F(1, 48^{ggadj}) = 1.36, p = 0.25$

Appendix 35: The Main Effects and Interactions for SDQ 3-Factor Structure Sub-scales

	Self-Report		
Interaction	**Extrovert Sub-scale**	**Introvert Sub-scale**	**Hyperactivity Sub-scale**
Treatment ^{x} Timepoint	$F(1, 48^{ggadj}) = 0.69, p = 0.41$	$F(1, 48) = 0.04, p = 0.85$	$F(1, 48^{ggadj}) = 0.19, p = 0.67$
Timepoint ^{x} Gender	$F(1, 48^{ggadj}) = 0.09, p = 0.77$	$F(1, 48) = 0.85, p = 0.36$	$F(1,48^{ggadj}) = 0.35, p = 0.56$
Timepoint ^{x} Age	$F(1, 48^{ggadj}) = 0.25, p = 0.62$	$F(1, 48) = 0.03, p = 0.86$	$F(1, 48^{ggadj}) = 0.08, p = 0.78$
Timepoint x IQ	$F(1, 48^{ggadj}) = 0.56, p = 0.46$	$F(1, 48) = 0.33, p = 0.57$	$F(1, 48^{ggadj}) = 0.00, p = 0.98$
Treatment ^{x} Gender	$F(1, 48^{ggadj}) = 0.27, p = 0.61$	$F(1, 48) = 0.99, p = 0.32$	$F(1, 48^{ggadj}) = 0.68, p = 0.41$
Treatment ^{x} Age	$F(1, 48^{ggadj}) = 3.24, p = 0.89$	$F(1, 48) = 0.17, p = 0.69$	$F(1, 48^{ggadj}) = 1.36, p = 0.25$
Treatment x IQ	**$F(1, 48^{ggadj}) = 9.41, p \leq 0.01$**	$F(1, 48) = 0.17, p = 0.69$	**$F(1, 48^{ggadj}) = 5.68, p \leq 0.05$**
Treatment ^{x} Timepoint ^{x} Gender	$F(1, 48^{ggadj}) = 2.69, p = 0.11$	$F(1, 48) = 0.01, p = 0.93$	$F(1, 48^{ggadj}) = 0.33, p = 0.57$
Main Effects			
Treatment	$F(1, 48^{ggadj}) = 0.30, p = 0.59.$	$F(1, 48) = 0.40, p = 0.53$	$F(1, 48^{ggadj}) = 3.26, p = 0.08$
Timepoint	$F(1, 48^{ggadj}) = 0.03, p = 0.86$	$F(1, 48) = 0.00, p = 0.96$	$F(1, 48^{ggadj}) = 0.05, p = 0.83$
Gender	**$F(1, 48^{ggadj}) = 10.10, p \leq 0.01$**	$F(1, 48) = 1.12, p = 0.30$	**$F(1, 48^{ggadj}) = 9.45, p \leq 0.01$**
Covariate Effects			
IQ	$F(1, 48^{ggadj}) = 0.17, p = 0.69$	$F(1, 48) = 0.49, p = 0.49$	$F(1, 48^{ggadj}) = 3.76, p = 0.06$
Age	$F(1, 48^{ggadj}) = 2.70, p = 0.11$	$F(1, 48) = 0.03, p = 0.87$	$F(1, 48^{ggadj}) = 0.36, p = 0.55$

Appendix 36: The Main Effects and Interactions for RCADS Total Score

	RCADS Total Score		
Interaction	***Parents Response (with outlier)***	***Parents Response (without outlier)***	***Self-Report Response***
Treatment ^{x} Timepoint	$F(1, 28^{ggadj}) = 0.67, p = 0.42$	$F(1, 26^{ggadj}) = 0.18, p = 0.68$	$F(1, 47^{ggadj}) = 0.30, p = 0.59$
Timepoint ^{x} Gender	$F(1, 28) = 0.11, p = 0.75$	$F(1, 26^{ggadj}) = 0.01, p = 0.94$	$F(1, 47^{ggadj}) = 0.46, p = 0.50$
Timepoint ^{x} Age	$F(1, 28^{ggadj}) = 3.42, p = 0.08$	$F(1, 25^{ggadj}) = 1.33, p = 0.26$	$F(1, 47^{ggadj}) = 0.30, p = 0.59$
Timepoint x IQ	$F(1, 28^{ggadj}) = 0.75, p = 0.39$	$F(1, 26^{ggadj}) = 0.33, p = 0.57$	$F(1, 47^{ggadj}) = 0.00, p = 0.99$
Treatment ^{x} Gender	$F(1, 28^{ggadj}) = 0.00, p = 0.98$	$F(1, 26^{ggadj}) = 0.02, p = 0.90$	$F(1, 47^{ggadj}) = 0.10, p = 0.75$
Treatment ^{x} Age	**$F(1, 28^{ggadj}) = 13.58, p \leq 0.001$**	$F(1, 24^{ggadj}) = 0.53, p = 0.47$	$(1, 47^{ggadj}) = 1.29, p = 0.26$
Treatment x IQ	$F(1, 28^{ggadj}) = 0.55, p = 0.46$	$F(1, 26^{ggadj}) = 0.01, p = 0.91$	$F(1, 47^{ggadj}) = 0.68, p = 0.41$
Treatment ^{x} Timepoint ^{x} Gender	$F(1, 28^{ggadj}) = 0.08, p = 0.76$	$F(1, 26^{ggadj}) = 0.00, p = 0.99$	$F(1, 47^{ggadj}) = 0.03, p = 0.85$
Main Effects			
Treatment	**$F(1, 28) = 14.17, p \leq 0.001$**	$F(1, 24^{ggadj}) = 0.47, p = 0.50$	$(1, 47^{ggadj}) = 0.49, p = 0.49$
Timepoint	$F(1, 28^{ggadj}) = 2.13, p = 0.16$	$F(1, 25^{ggadj}) = 0.99, p = 0.33$	$F(1, 47^{ggadj}) = 0.22, p = 0.64$
Gender	$F(1, 28^{ggadj}) = 1.02, p = 0.32.$	$F(1, 26^{ggadj}) = 0.78, p = 0.39$	$F(1, 47^{ggadj}) = 1.14, p = 0.29$
Covariate Effects			
IQ	$F(1, 28^{ggadj}) = 3.44, p = 0.07$	$F(1, 26^{ggadj}) = 1.26, p = 0.88$	$F(1, 48^{ggadj}) = 3.69, p = 0.06$
Age	**$F(1, 2^{ggadj}) = 6.54, p \leq 0.05$**	$F(1, 25^{ggadj}) = 0.02, p = 0.88$	$F(1, 48^{ggadj}) = 0.02, p = 0.88$

Appendix 37: The Main Effects and Interactions for SNAP-IV Sub-scale Scores

	SNAP-IV Inattention Sub-scale		**SNAP-IV Hyperactivity Sub-scale**	
Interaction	***Teacher Response***	***Parents Response***	***Teacher Response***	***Parent Response***
Treatment ^{x} Timepoint	$F(1, 31^{ggadj}) = 1.47, p = 0.24$	$F(1, 17^{ggadj}) = 0.96, p = 0.34$	$F(1, 31^{ggadj}) = 0.06, p = 0.81$	$F(1, 21^{ggadj}) = 0.36, p = 0.55$
Timepoint ^{x} Gender	$F(1, 31^{ggadj}) = 0.000, p = 1.00$	$F(1, 18^{ggadj}) = 0.93, p = 0.35$	$F(1, 31^{ggadj}) = 2.08, p = 0.16$	$F(1, 22^{ggadj}) = 0.10, p = 0.76$
Timepoint ^{x} Age	$F(1, 32^{ggadj}) = 1.56, p = 0.22$	$F(1, 19^{ggadj}) = 0.56, p = 0.46$	$F(1, 32^{ggadj}) = 0.13, p = 0.72$	$F(1, 22^{ggadj}) = 0.09, p = 0.76$
Timepoint x IQ	$F(1, 32^{ggadj}) = 0.64, p = 0.43$	$F(1, 18^{ggadj}) = 0.00, p = 1.00$	$F(1, 32^{ggadj}) = 0.09, p = 0.76$	$F(1, 22^{ggadj}) = 0.06, p = 0.81$
Treatment ^{x} Gender	$F(1, 31^{ggadj}) = 0.85, p = 0.36$	$F(1, 19^{ggadj}) = 1.95, p = 0.18$	**$F(1, 31^{ggadj}) = 4.72, p < 0.05$**	$F(1, 22^{ggadj}) = 0.18, p = 0.68$
Treatment ^{x} Age	$F(1, 18^{ggadj}) = 0.64, p = 0.43$	$F(1, 16^{ggadj}) = 2.86, p = 0.11$	$F(1, 18^{ggadj}) = 1.58, p = 0.23$	$F(1, 22^{ggadj}) = 2.08, p = 0.16$
Treatment x IQ	$F(1, 31^{ggadj}) = 2.43, p = 0.13$	$F(1, 15^{ggadj}) = 0.03, p = 0.87$	$F(1, 32^{ggadj}) = 1.64, p = 0.21$	$F(1, 21^{ggadj}) = 5.01, p = 0.04$
Treatment ^{x} Timepoint ^{x} Gender	$F(1, 31^{ggadj}) = 0.11, p = 0.74$	$F(1, 19^{ggadj}) = 0.20, p = 0.66$	$F(1, 31^{ggadj}) = 0.02, p = 0.90$	$F(1, 22^{ggadj}) = 0.14, p = 0.72$
Main Effects				
Treatment	$F(1, 18^{ggadj}) = 2.03, p = 0.17$	$F(1, 16^{ggadj}) = 2.54, p = 0.13$	$F(1, 18^{ggadj}) = 2.66, p = 0.12$	$F(1, 22^{ggadj}) = 0.34, p = 0.47$
Timepoint	$F(1, 32^{ggadj}) = 0.85, p = 0.36$	$F(1, 19^{ggadj}) = 0.37, p = 0.55$	$F(1, 32^{ggadj}) = 0.29, p = 0.60$	$F(1, 22^{ggadj}) = 0.02, p = 0.90$
Gender	**$F(1, 31^{ggadj}) = 17.43, p < 0.001$**	**$F(1, 20^{ggadj}) = 11.01, p \leq 0.01$**	**$F(1, 31^{ggadj}) = 17.64, p < 0.001$**	$F(1, 22^{ggadj}) = 1.94, p = 0.18$
Covariate Effects				
IQ	$F(1, 31^{ggadj}) = 1.17, p = 0.29$	$F(1, 20^{ggadj}) = 2.77, p = 0.11$	$F(1, 31^{ggadj}) = 0.49, p = 0.49$	$F(1, 22^{ggadj}) = 1.24, p = 0.28$
Age	**$F(1, 29^{ggadj}) = 12.78, p \leq 0.001$**	$F(1, 20^{ggadj}) = 0.10, p = 0.76$	**$F(1, 29^{ggadj}) = 17.64, p \leq 0.01$**	$F(1, 22^{ggadj}) = 0.08, p = 0.78$