a) Title

Project title: Chronic and acute effects of artificial colourings and preservatives on children's behaviour.

Project code: T07040

Dates: 1 September 2004 to 28 February 2007

Contractor: School of Psychology, University of Southampton, Highfield, SOUTHAMPTON SO17 1BJ

PI: Prof. Jim Stevenson

Co-PIs Prof. Edmund Sonuga-Barke
Prof. John Warner

Report date: 18 June 2007 - Revised 10 September 2007

'This research project was commissioned by the Food Standards Agency. The contents of this final technical report represent the views of the researchers and not necessarily those of the Food Standards Agency.'

b) Executive Summary

This Technical Report is structured such that the main text provides an integrative overview of the results of the studies on 3 year old and on 8 to 9 year old children. Annexes 1 and 2 contain the detailed accounts of the design, sampling, measurement and data analytic methods and findings for the study of 3 year old and 8/9 year olds respectively.

There is a longstanding suggestion that artificial food colours and other food additives such as preservatives (AFCA) influence behaviour in children. It is over 30 years since Feingold made his initial claims of the detrimental effect of AFCA on children's behaviour. [1] The main putative effect of AFCA is to produce overactive, impulsive and inattentive behaviour, i.e. hyperactivity, which is a pattern of behaviour that shows substantial individual differences in the general population. Children who show this behaviour pattern to a marked degree are likely to be diagnosed with attention deficit hyperactivity disorder (ADHD).[2] Despite the failure of early studies to identify the range of proposed adverse affects,[3] a more recent meta-analysis of double-blind, placebo-controlled trials has shown a significant effect of AFCA on the behaviour of children with ADHD.[4] The possible benefit in reducing the level of hyperactivity of the general population by the removal of AFCA from the diet is less well established. There is some evidence from our earlier study on the Isle of Wight of adverse effects on hyperactivity measured by parental ratings for 3 year old children of one mix of additives.[5] These findings required replication on three year old children and to establish if the effects could be found using a wider range of measures of hyperactivity. The present community based double-blinded, placebo-controlled food challenge (DBPCFC) was designed to extend the age range studied to include 8/9 year old children to determine if the effects could also be detected in middle childhood.

Main results

The material in this technical report has been incorporated into one paper that has been published in the Lancet.[6] The evidence we have obtained is that certain mixtures of artificial colours and sodium benzoate preservative (referred to in this report as Mix A and Mix B) adversely affect the hyperactive behaviour of children in some age groups compared with a placebo.

The results replicate and extend the findings from our earlier study.[5] The specific findings were that with Mix A there was a significant ($p < 0.05$) adverse effect on the average hyperactive behaviour of 3 year old children as measured using the chosen outcome measure of a Global Hyperactivity Aggregate (GHA) and based on the primary analysis of the study (the whole cohort). In contrast, Mix B was without significant effect on the behaviour of 3 year old children. The reverse picture was seen with 8/9 year old children. In this case, compared with placebo, Mix B had a significant adverse effect on the behaviour of children ($p<.05$). However, for the whole cohort Mix A was without significant effect on the behaviour of 8/9 year old children. While an aggregate score (GHA) was the primary outcome measure for this study it is

noted that, as in the previous Istudy carried out on the Isle of Wight, that the parental ratings of behaviour are a significant contributor to the GHA score.

The importance of these findings is that they confirm that the adverse effect of certain artificial food colours that has been implicated in children with hyperactive syndromes[4] can also be demonstrated in two samples taken from the general population.

For both 3 and 8-9 year olds a range of factors were examined to determine if they made the child more vulnerable to the effects of the food colour and benzoate preservative mixes. None of the social factors examined (age, gender, pre-trial diet, maternal education) moderated the effects of the active mixes at either age. However, for the 3 year old children consuming more than 85% of the drink challenges a polymorphisms in the histamine N-methyltransferase gene (HNMT Thr105Ile) moderated the effect on the GHA of Mix A compared to Placebo. Specifically the absence of HNMT 105Ile in the genotype made the 3 year-old children more vulnerable to the adverse effects of the Mix A additive mixture. For the 8 to 9 year old children only, this same moderating effect of the absence of HNMT 105Ile in the genotype was found for Mix B. In addition for this the HNMT T939C polymorphism (specifically the absence of the 939C allele in the genotype) made the 8 to 9 year old children more vulnerable to the adverse effects of the Mix A and B additive mixtures.

Conclusions

This study has shown that effects on behaviour may be associated with intake of some mixtures of specific food colours and the preservative sodium benzoate. Two mixtures were examined in unselected populations of 3 and 8/9 year olds. One mixture (Mix A) was shown to cause a significant increase in hyperactivity in one age group (3 year olds), while the second mixture (Mix B) caused significant effects in the other age group (8/9 year olds).

The size of the effects of the colour and preservative mixtures studied on the average hyperactivity score is lower than that reported for clinical samples. We recognise that hyperactivity is a behaviour influenced by a wide range of experiential and biological factors. It is known that there are major genetic influences on hyperactivity[7] and this study has shown additionally that, when only children consuming more than 85% of the challenge drinks are considered, differential sensitivity to the mixture of food colours and preservative resulting from certain genetic polymorphisms is one route by which genetic influences on hyperactivity may be mediated. Although the results of the study suggest that some mixtures of certain food colours and benzoate preservative may effect the level of hyperactive behaviour in children, removal of these additives would not be a panacea for ADHD.

c) Glossary

HNMT : histamine N-methyltransferase gene
ADHD : attention deficit hyperactivity disorder
AFCA : artificial food colours and other food additives
GHA : global hyperactivity aggregate
DBPCFC ; double blind placebo controlled food challenge.
ES : effect size (mean on mix minus mean on placebo / SD on placebo)
HI : hyperactivity index (dependent variable used in acute challenge study)
DRD4: dopamine D4 receptor gene
DAT1: dopamine transporter gene
ADRA2A : adrenergic receptor alpha 2A gene

d) Aims and objectives of the Investigation

There is a longstanding suggestion that artificial food colours and preservatives influence behaviour in children. It is over 30 years since Feingold made his initial claims of the detrimental effect of AFCA on children's behaviour [1]. The main putative effect of AFCA that has been proposed is to produce overactive, impulsive and inattentive behaviour, i.e. hyperactivity, which is a pattern of behaviour that shows substantial individual differences in the general population. Children who show this behaviour pattern to a marked degree are likely to be diagnosed with attention deficit hyperactivity disorder (ADHD)[8]. A meta-analysis of double-blind, placebo-controlled trials has indicated a significant effect of AFCA on the behaviour of children with ADHD.[4] Whether AFCA's have a similar effect in the general population has not been conclusively demonstrated. There is some evidence from our earlier study on the Isle of Wight of adverse effects on hyperactivity measured by parental ratings for 3 year old children in response to one mix of additives (four artificial colours and sodium benzoate preservative), but these findings were not replicated by the behavioural assessments conducted in a more controlled setting. [5]

The present study was undertaken to replicate and extend previous research on the Isle of Wight. The study was based on a sample of children selected from the general population. This included both 3-year olds as in the previous study but also 8-9 year olds, to test whether effects could be identified in an older sample. As well as replicating a test of the particular mix of artificial food colours and preservative we used previously (Mix A - sunset yellow (E110), carmoisine (E122), tartrazine (E102), ponceau 4R(E124) and sodium benzoate (E211)), we wanted to establish whether giving children a mixture of colours and preservative more representative of levels and types of these additives experienced by children at the time of commissioning this study (2004), would produce effects on behaviour (Mix B - sunset yellow (E110), carmoisine (E122), quinoline yellow (E104), allura red AC(E129)) and sodium benzoate(E211)). In addition to determining whether these mixtures of additives would have an effect on children's behaviour across the general population, it was possible that specific children would be more sensitive to the effects. In particular we tested whether a number of genetic polymorphisms made individual children more or less sensitive to any effects of the mixtures. In addition,we wanted to test in a "proof of principle" study whether behavioural and metabolic changes in children were apparent when given a daily dose of a mixture of food colours and benzoate preservative in an acute challenge.

The main study was designed to test the following primary hypothesis:

> Mixtures of certain artificial colourings and a benzoate preservative increase the mean level of hyperactive behaviour in children from the general population.

The secondary research questions were:

> Is this response seen in teacher ratings, direct observations of behaviour and test performance as well as in parent ratings?

Is the response to food additives and colourings related to initial levels of ADHD i.e. is the response greater in children at the extreme end of the continuum?

Do genetic differences moderate the effect?

What are the metabolic and neuropsychological mediators of any effect?

The study also examined whether the effects of the food colourings and the preservative were more clearly seen for those children consuming an adequate amount of the total challenge. This was set at 85% or more of the drinks presented to the children. A complete case analysis was also conducted to check whether the treatment of missing data in the analysis had introduced any bias.

e) Experimental Procedures

Participants

Figures 1 and 2 present details of recruitment and participation in the study, for 3-year-old and 8/9-year-old children, respectively. The study sample was drawn from a population of children aged between 3 years and 4 years, 2 months, registered in early-years settings (nurseries, day nurseries, preschool groups, playgroups) and from children aged between 8 and 9 years attending schools in Southampton, UK. To ensure that the study sample included children from the full range of socioeconomic backgrounds, schools were recruited based on the number of children receiving free school meals (an index of social disadvantage). The distribution of the percentage of children receiving free meals in the schools taking part indicated the proportions for the city as a whole. To further check on how representative the sample was, teachers completed a hyperactivity questionnaire[9] for all 3-year-old and 8/9-year-old children.

Parents who returned an expression of interest form were contacted by phone and a home visit arranged. On this visit, a research assistant and the study dietitian, provided full information about the study and its dietary implications, and written informed consent was obtained. The study dietitian also obtained a report based on 24-h recall by the parent of the child's pretrial diet, which allowed an assessment of baseline levels of the number of foods containing additives consumed by the child in the previous 24 h. The study was approved by the local research ethics committee (reference no 04/Q1702/61) and written informed consent was obtained from parents. Participating early-years settings received £250 and each school £500 as a contribution towards school funds for the benefit of all children.

Figure 1 Enlistment of 3 year old participants

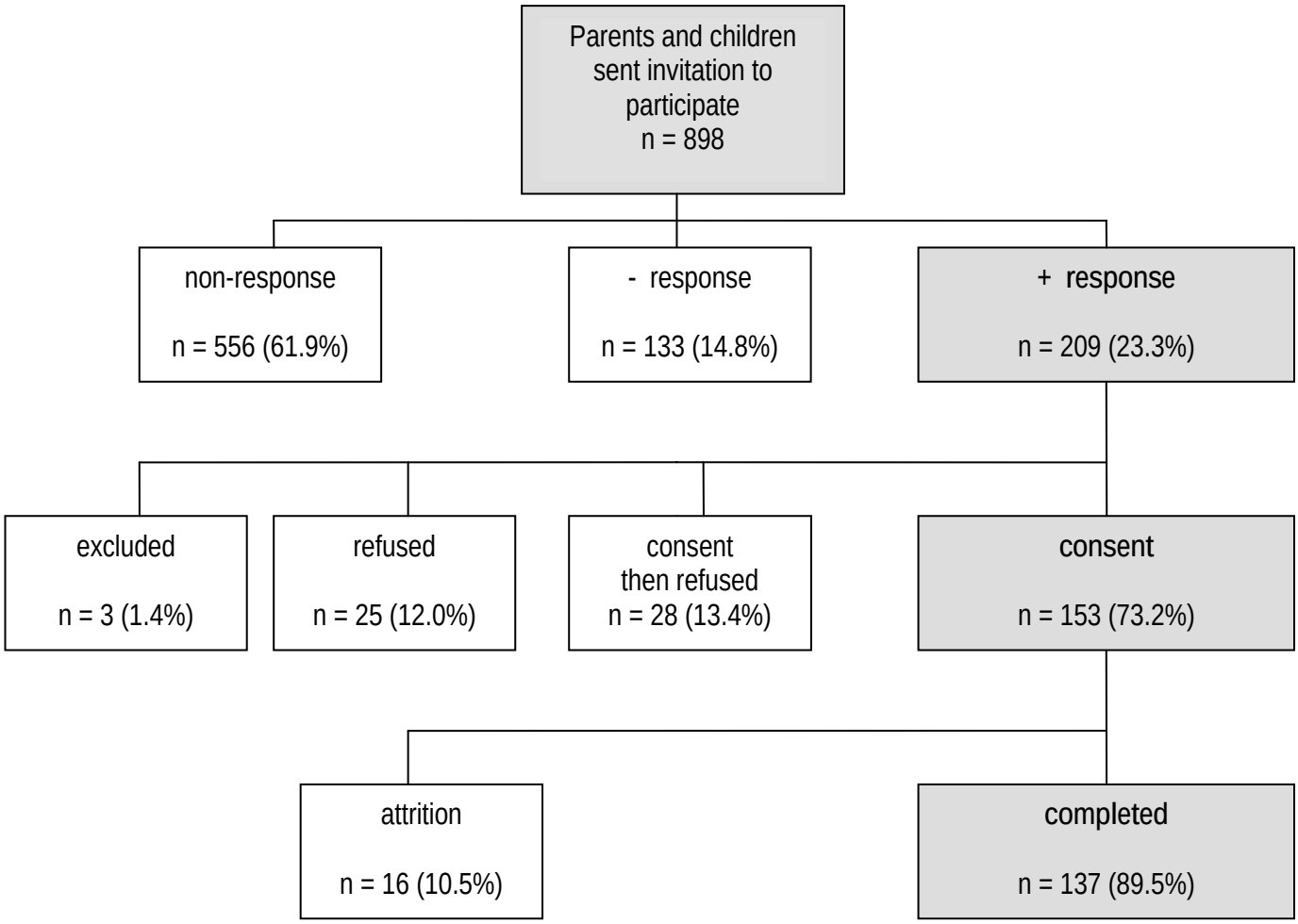
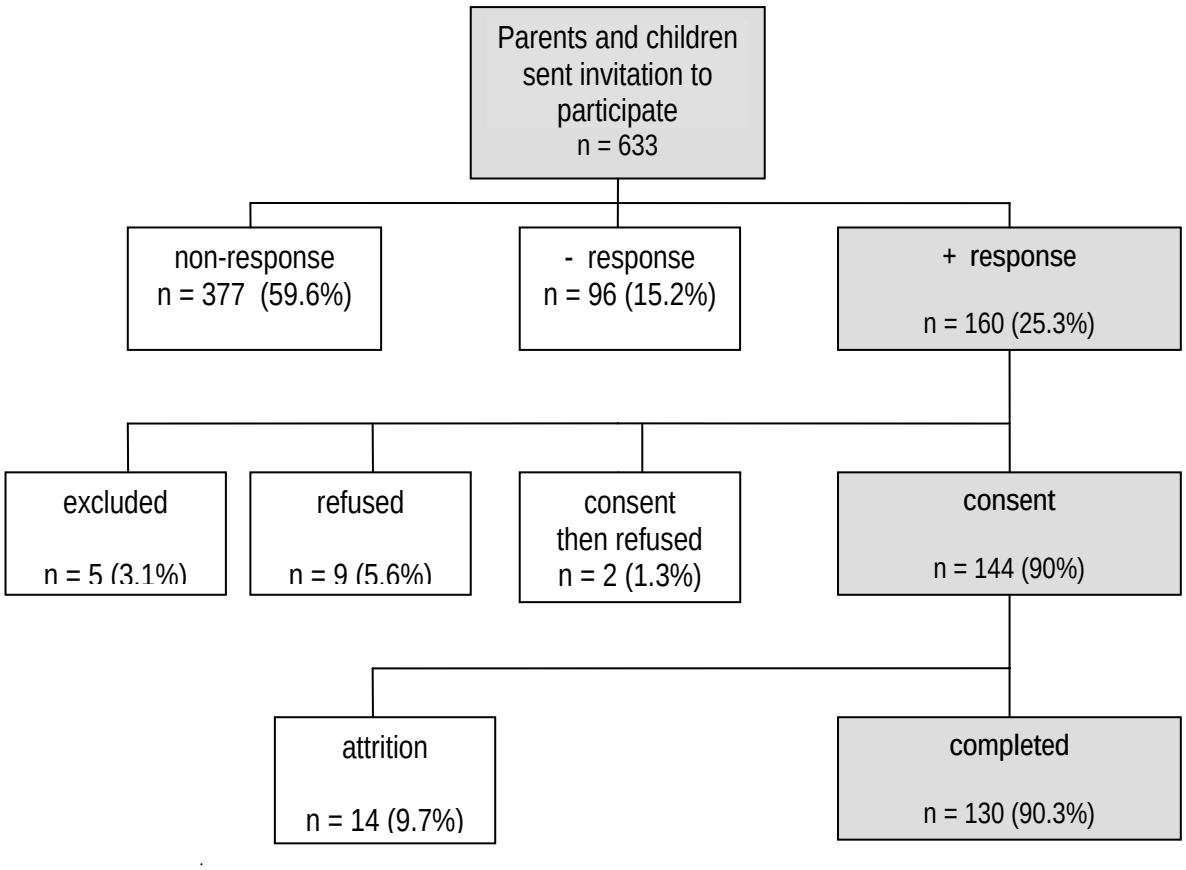


Figure 2 Enlistment of 8/9 year old participants



Study design and challenge protocols

The study design and challenge protocols for both ages were similar. Children were entered into this study with a within-subject crossover between two active Mixes (A and B) and a placebo drink. The two active mixes differed both in the quantities of additives and the specific additives included. Mix A was similar to the active challenge used in the Isle of Wight study,[5] and Mix B was selected to indicate the current average daily consumption of food additives by 3-year-old and 8/9-year-old children in the UK.[10] Both mixes included sodium benzoate, which had been included in the challenge on the Isle of Wight study and in previous studies.[11,12]

Mix A for 3-year-old children included 20 mg of artificial food colourings (5 mg sunset yellow [E110], 2·5 mg carmoisine [E122], 7·5 mg tartrazine [E102], and 5 mg ponceau 4R [E124, Forrester Wood, Oldham, UK]) and 45 mg of sodium benzoate [E211, Sigma Aldridge, Gillingham, UK]). Active Mix B included 30 mg of artificial food colourings (7·5 mg sunset yellow, 7·5 mg carmoisine, 7·5 mg quinoline yellow [E104], and 7·5 mg allura red AC [E129]) and 45 mg of sodium benzoate.

Mix A amounts for 8/9-year-old children were multiplied by 1·25 to account for the increased amount of food consumed by children at this age. Therefore, Mix A included 24·98 mg of artificial food colourings (6·25 mg sunset yellow, 3·12 mg carmoisine, 9·36 mg tartrazine, and 6·25 mg ponceau 4R) and 45 mg of sodium benzoate. Active Mix B included 62·4 mg of artificial food colourings (15·6 mg sunset yellow, 15·6 mg carmoisine, 15·6 mg quinoline yellow, and 15·6 mg allura red AC) and 45 mg of sodium benzoate.

After a week on their typical diet (week 0: baseline diet), the artificial colours to be used in the challenges and sodium benzoate were withdrawn from their diet for 6 weeks. Over this period when challenge with active or placebo drinks were given, additive withdrawal continued (week 1: withdrawal period but receiving placebo; weeks 2, 4, and 6: challenge with randomisation to two active periods and one placebo period; weeks 3 and 5: washout continuing on placebo). During this period, 3-year-old children received the challenge and washout-placebo drinks on a weekly basis and consumed mixed fruit juices (placebo or active) at home (300 mL/day for 3-year-old children, 625 mL/day for 8/9-year-old children), provided in identical sealed bottles. At the beginning of the study, children were assigned by the study administrator by a random-number generator to receive one of six possible sequences of placebo, active Mix A, or active Mix B challenges across weeks 2, 4, and 6.

A masked testing by two independent panels of 20 young adults showed that the active and placebo juice drinks could not be differentiated. When asked if the mix contained additive, 16 (40%), 21 (52%), and 26 (65%) adults responded positively for Mix A, Mix B, and placebo, respectively. We recorded

no significant differences between these proportions (Friedman test, χ^2=4·412, df=2). Therefore, no reliable differences were seen between the look and taste of the drinks. Such differences as there were appeared to be in the direction of the placebo drink being seen as more likely to contain additives i.e. which is the direction of being conservative to our hypothesised effects.

The only difference in the composition of the placebo and active mixes was the presence of the AFCA in the active mix with some variation in the proportions of the fruit juices to ensure matching colour and taste for the placebo and active drinks. The child's family and the research team were masked to the challenge allocation. The study administrator assigned the challenge sequence and assisted in the preparation and packaging of juice drinks that were then delivered by the masked research team to homes every week, when questionnaires and other forms were obtained and dispensed. Parents completed a daily diary of juice consumption and compliance with the diet over the study period. Parents also recorded a mistake event when a child consumed a portion of food containing the artificial colours or sodium benzoate. Any bottles containing juice not consumed in the previous week were obtained, returned to the study office, and measured to help validate, if possible, parental reports of juice consumption by children.

Global hyperactivity aggregate (GHA)

Three measures of behaviour were used to calculate GHA for 3-year-old children, with an additional measure for 8/9-year-old children. First, the abbreviated ADHD rating scale IV (teacher version)[9] was used. A total score was obtained for ten of the 18 items (inattentive=5, hyperactive=5) in this questionnaire, which was completed to describe the frequency of the specific behaviours displayed over the past week, for every week of the study. Parent behaviour was the second measure, by use of the abbreviated Weiss-Werry-Peters (WWP) hyperactivity scale,[13] which has been used in several studies to assess hyperactivity.[14,15] Interparent agreement is good for ratings of childhood behaviour (r=0·82).[16] Parents rated their child's behaviour during the previous week for seven items previously used (switching activities; interrupting or talking too much; wriggling; fiddling with objects or own body; restless; always on the go; concentration),[5] from which we obtained a total score. For 8/9-year-old children, we used an abbreviated ADHD rating scale IV (parent version)[17] to measure parent behaviour, whereby a ten-item questionnaire was completed by parents every week.

A third measure was the classroom observation code,[18] which assesses the occurrence of 12 mutually exclusive behaviours during structured didactic teaching and during periods of independent work under teacher supervision. To develop this measure, the behaviours had been selected to indicate components of ADHD that are shown in the classroom. After observers (psychology graduates) were given extensive training, the inter-rater reliability of the classroom observation code, tested before and during the study, exceeded 0·87. Children were observed for 24 min every week (three

observation sessions of 8 min each) and a total weekly mean score was derived from the total score over every session. The code was slightly modified for 3-year-old children, since preschool children in the UK are not usually given structured or didactic teaching sessions and tend to engage in activities rather than in tasks. Observation took place over a range of activities and the off-task category in the code was scored when the child switched activities.

A fourth measure for 8/9-year-old children was the Conners' continuous performance test II (CPTII),[19] a test using visual stimuli of 14-min duration and is widely used to assess attention and the response inhibition component of executive control. We used four scores (SE of reaction time, % of commission errors, d´ [discriminability index], and β) to derive a weekly aggregate score. This subset of indicators from the CPTII has been shown to be highly correlated with the ADHD rating scale.[20]

The GHA was developed to measure individual differences in hyperactivity using different sources (teacherratings, parent ratings, direct observation, and a computerised test) and covering the components of hyperactivity (overactivity, impulsivity, and inattention). Weekly scores for every child were standardised to time 0 at baseline (T0). Weekly standardised (z) aggregate scores were calculated as: (score minus mean score at T0) divided by SD at T0. The GHA was an equally weighted aggregate of the weekly z-scores, and calculated only when at least two (or three for 8/9-year-old children) of these behaviour scores were present for any week (one of which being for the classroom observation code) and averaged across the number of available scores. A high GHA indicates more hyperactivity.

Statistical analysis

Although the study designs for the two age groups were similar, the difference in composition of the GHA, and in the dose of AFCA used, meant that data from the two studies could not be analysed jointly. Therefore, we treated the studies as parallel but independent.

Linear mixed-model methods[21,22] in SPSS (version 14.0) were used to analyse data. Several possible covariates were thought to be significantly related to GHA (eg, sex). Two models were tested separately for each age for the effects on GHA in challenge weeks. Model 1 used the challenge type alone as a fixed effect testing for Mix A against placebo and Mix B against placebo. In Model 2, in addition to challenge type, the effects of the following factors were adjusted for: week during study, sex, GHA in baseline week, number of additives in pretrial diet, maternal educational level, and social class. A compound symmetry covariance matrix provided the best-fit model for 3-year-old children and an unstructured covariance matrix for 8/9-year-old children. The study was powered to detect differences between the active and placebo periods and, accordingly in each case, the effects of Mix A and Mix B were compared with that of placebo. We anticipated that the additional controls on placebo effects would result in an effect size smaller than that achieved in the Isle of Wight study.[5] A sample of 80 children had 80% power

at α=0·05 to identify an effect size of 0·32—ie, the magnitude of the difference in GHA mean score changes (SD). This value was lower than that achieved in the previous study (0·51). We were uncertain about the number of children and families who would comply with the demands of a 7-week study, so we set a target of 120 children to reduce the effect of attrition on power, which was eventually exceeded in both age groups.

The analyses were replicated for the full sample, a high consumption group (≥85% consumption of drinks in any challenge week), and a complete case group (≥85% consumption in all challenge weeks and no missing GHA). The high consumption and complete case groups were included to determine whether non-compliance and the method of handling missing data affected the pattern of results. Analysis was per-protocol.

This clinical trial is registered with Current Controlled Trials (registration number ISRCTN74481308).

For details of the methods of the acute "proof of principle" study, please see p.28 and Annex 2 Section 22

Further details about the methodology for the study of 3 year olds can be found in Annexe 1 Sections 2 to 9. Similarly for the study of 8-9 year olds can be found in Sections 2 to 9 of Annex 2.

The analyses summarised here are based on the primary outcome measure for the study, which was an aggregate measure of hyperactvity incorporating both subjective and objective measures of assessment, labelled the GHA (global hyperactivity aggregate). The GHA is the main outcome and is based on the aggregated z-scores of observed behaviours and ratings by teachers and parents, plus, for 8/9 year old children, a computerised test of attention. Analyses of the behavioural scores in response to additive challenge versus placebo obtained from each individual behavioural assessment measure separately, were performed as a secondary outcome measure and are presented in Sections 21 of both Annexes 1 and 2 and are summarised in the Secondary Research Questions section below.

f) Results

Primary research question

3 year olds

Study sample

Of 153 children (mean age 43.5 months, SD: 4.5, range: 35 to 53) enlisted to the study, 79 were boys (mean age 43.5 months, SD: 4.6, range: 35 to 52) and 74 girls (mean age 43.4 months, SD: 4.3, range: 35 to 53). Table 1 provides characteristics separately for the the primary outcome, i.e. the whole sample (n=153), a sub-group of children (n=132) with what we defined upon reviewing the results as an acceptable (≥85%) consumption of drinks, and a 'complete case' group with ≥85% consumption and no missing GHA data (n=73). No significant differences in these background characteristics were found between sub-samples or between groups of children assigned to receive the challenge drinks in different orders over each of the 6 periods (weeks).

Consumption of drinks

A few children dropped out of the study over its duration, for various reasons as detailed in Section 3 of Annex 1. Of those children who dropped out of the study, 12 had a mean of 41% consumption of drinks in the first challenge week and data was missing for 4 children. Of the children who completed the study, 128/137 (93%) consumed more than two thirds of all drinks with 80% of these consuming ≥85% (at least 6 out of 7 daily drinks per week). Only 1 of the remaining 9 children drank less than 50% of placebo and active drinks over the study period.

The occurrence of dietary infractions or 'mistakes' by children over the study period was low (33% = 0; 31% = 1 to 2; 18.3% = 3 to 4; 17% > 4) (see Annex 1 section 19 for more details). The rate of infractions did not differ during active and placebo weeks.

Effects of Challenge

Of 153 children, 117 (76.5%) had full GHA data over active and placebo weeks, 19 (12.4%) had 2 GHA scores and 1 child had 1 GHA. Children who left the study (n=16) provided 1 GHA score each (n=12) or had missing data (n=4).

In order to identify potential moderating or confounding effects of a number of variables on the behavioural response of the 3 year old children to the challenge mixtures, we examined their relationship to baseline GHA (Table 2). Baseline GHA was significantly related to GHA at all subsequent time points. When baseline GHA was included in the mixed models analyses, the effect of gender on behaviour became non-significant. This arose because boys had higher hyperactivity baseline scores than girls (a widely reported finding). The putative effects of gender were more appropriately seen to be reflected in higher baseline scores than to be an effect of gender per se. Preliminary

analyses showed that there was no effect of time (week) on the GHA and no carryover effect from the previous challenge week to the next challenge week on the GHA score.

Table 3 shows the unadjusted mean scores for the GHA for the three challenges.

The test of the effect of challenge (Mix A or Mix B versus placebo) on GHA (Model 1) can be found in Table 4 and for the same analysis repeated but including a number of potential confounding factors, including the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2). There were no significant interactions between challenge and these variables. Analyses are shown separately for the full sample (the primary outcome), and *post-hoc* analyses of two sub groups who had high consumption (case included if $\geq$85% consumption of drinks in any challenge week) and complete case ($\geq$85% consumption in all challenge weeks and no missing GHA). The estimated marginal means in standard deviation units of the GHA at baseline (i.e. combined Z scores) for the challenges adjusted for the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class (Model 2) are displayed in Figure 3. The analysis of the whole sample with the controls incorporated in Model 2 showed a significant effect of Mix A on GHA compared to placebo (effect size = 0.20; $CI_{.95}$ .01 to .39; p<.05). The effect of Mix B did not reach significance (effect size = 0.17; $CI_{.95}$ -.03 to .36; ns).

When the Model 2 analyses are restricted to those children with $\geq$85% juice consumption (see Table 4), the adverse effect of Mix A on behaviour remains statistically significant (effect size = 0.28; $CI_{.95}$ .05 to .51; p<.05) but is non-significant for Mix B(effect size = 0.19; $CI_{.95}$ -.04 to .41; ns) (p=.074). The complete case group shows the same pattern of results, i.e. a statistically significant effect of Mix A (effect size = 0.32; $CI_{.95}$ .05 to .60; p<.05) but not of Mix B (effect size = 0.21; $CI_{.95}$ -.06 to .48; ns).

Table 1: Characteristics of parents of 3 and of 8/9 year old children

	3 year olds			8/9 year olds		
	Children in total sample analysis n =153 n(%)	Children in ≥85% drunk analysis n = 133 n(%)	Children in complete case analysis n = 73 n(%)	Children in total sample analysis n =144 n (%)	Children in ≥85% drunk analysis n = 119 n (%)	Children in complete case analysis n = 91 n (%)
Racial background						
White	126 (82.4)	117 (88.0)	67 (91.8)	130 (90.3)	110 (92.4)	85 (93.4)
Other	15 (9.8)	12 (9.0)	6 (8.2)	14 (9.7)	9 (7.6)	6 (6.6)
Missing data	12 (7.8)	4 (3.0)	-	-	-	-
Marital status						
Married/partner	127 (83.0)	110 (82.7)	60 (82.2)	115 (79.9)	99 (83.2)	80 (87.9)
Single/separated/divorced/widowed	26 (17.0)	23 (17.3)	13 (17.8)	29 (20.1)	20 (16.8)	11 (12.1)
NSSC* (Father)						
Higher occupations	34 (22.2)	30 (22.0)	15 (20.5)	37 (25.7)	35 (29.4)	29 (31.9)
Intermediate occupations	26 (17.0)	24 (18.2)	9 (12.3)	18 (12.5)	17 (14.3)	14 (15.4)
Lower occupations	51 (33.3)	48 (36.4)	32 (43.8)	44 (30.6)	39 (32.8)	30 (33.0)
Never worked/long term unemployed	4 (2.6)	4 (3.0)	4 (5.5)	7 (4.9)	7 (5.9)	6 (6.6)
No father present	26 (17.0)	23 (17.4)	13 (17.8)	29 (20.1)	20 (16.8)	11 (12.1)
Missing data	12 (7.8)	4 (3.0)	-	9 (6.3)	1 (0.8)	1 (1.1)
NSSC(Mother)						
Higher occupations	31 (20.3)	25 (18.8)	13 (17.8)	38 (26.4)	35 (29.4)	27 (29.7)
Intermediate occupations	18 (11.8)	18 (13.5)	12 (16.4)	26 (18.1)	25 (21.0)	20 (22.0)
Lower occupations	66 (43.1)	62 (46.6)	37 (50.7)	32 (22.2)	29 (24.4)	25 (27.5)
Never worked/long term unemployed	26 (17.0)	24 (18.0)	11 (15.1)	32 (22.2)	28 (23.5)	19 (20.9)
Missing data	12 (7.8)	4 (3.0)	-	16 (11.1)	2 (1.7)	-
Mother's education						
School to 16 (no qualifications / certificates below 'A' level)	53 (34.6)	50 (37.6)	23 (31.5)	60 (41.7)	54 (45.3)	40 (44.0)
'A' levels	61 (39.9)	55 (41.4)	35 (47.9)	42 (29.2)	39 (32.8)	32 (35.2)
University Degree/Post-graduate qualification	27 (17.7)	24 (18.0)	15 (20.5)	27 (18.8)	24 (20.2)	19 (20.9)
Missing data	12 (7.8)	4 (3.0)	-	15 (10.4)	2 (1.7)	-

*NSSC: National Statistics Social Class.[8]

Table 2: Effect of possible moderators on Baseline GHA for 3-year-olds (total N= 153)

Variable	Groups Mean (sd)	Mean (sd)	Test
Age Gp	Age ≤43m (n=66) -.14 (.69)	Age >43m (n=67) .07 (.66)	F(1,131)=3.18, p=.077
Gender	Male (n=67) .14 (.78)	Female (n=66) -.22 (.51)	F(1,131)=9.73, p=.002
Maternal education level	O levels or less (n=50) -.02 (.63)	A levels or higher (n=79) -.08 (.71)	F(1,127)=.215, p=.644
Pre-study diet	Pre-study diet Analysed as a continuous covariate		F(1,129)=4.00, p=.048

Table 3 Mean GHA for 3 year olds by challenge type

	Mix A		**Mix B**		**Placebo**	
	n=	**mean (SD)**	**n=**	**mean (SD)**	**n=**	**mean (SD)**
3 year olds						
Whole sample n=140	131	-.11 (1.03)	134	-.14 (1.03)	129	-.32 (1.11)
≥85% consumption n=130	104	-.11 (1.03)	108	-.15 (1.07)	99	-.39 (1.07)
Complete case n=73	73	-.14 (1.04)	73	-.26 (1.05)	73	-.44 (0.98)

Table 4. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=140	**≥85% consumption** n=130	**Complete case** n=73
Model 1			
Intercept	-.31 [-.49 to -.13] **	-.33 [-.53 to -.13] ***	-.44 [-.68 to -.21] ***
Challenge type			
Mix A –v- placebo	.20 [0.01 to .40] *	.24 [.02 to .47] *	.31 [.04 to .58] *
Mix B –v- placebo	.16 [-.04 to .35]	.16 [-.07 to .38]	.19 [-.08 to .46]
Model 2			
Intercept	-.54 [-.89 to -.18] **	-.51 [-.92 to -.11]	-.58 [-1.08 to -.09]*
Challenge type			
Mix A –v- placebo	.20 [.01 to .39] *	.28 [.05 to .51] *	.32 [.05 to .60] *
Mix B –v- placebo	.17 [-.03 to .36]	.19 [-.04 to .41]	.21 [-.06 to .48]
Week of study			
Wk 2 –v- Wk 6	.15 [-.05 to .34]	.15 [-.08 to .38]	.19 [-.08 to .46]
Wk 4 –v- Wk 6	.17 [-.03 to .36]	.23 [.00 to .46] *	.19 [-.09 to .46]
Gender	.18 [-.10 to .45]	.22 [-.07 to .51]	.05 [-.31 to .40]
Baseline GHA score	.46 [.26 to .66] ***	.54 [.31 to .76] ***	.36 [.06 to .66] *
Pre-trial diet	.08 [-.02 to .19]	.07 [-.04 to .18]	.09 [-.04 to .23]
Maternal education level	-.01 [-.29 to .28]	-.04 [-.34 to .26]	-.03 [-.41 to .35]
Maternal social class	.15 [-.44 to .13]	-.23 [-.53 to .08]	-.21 [-.58 to .16]

* p<.05, ** p<.01, *** p<.001

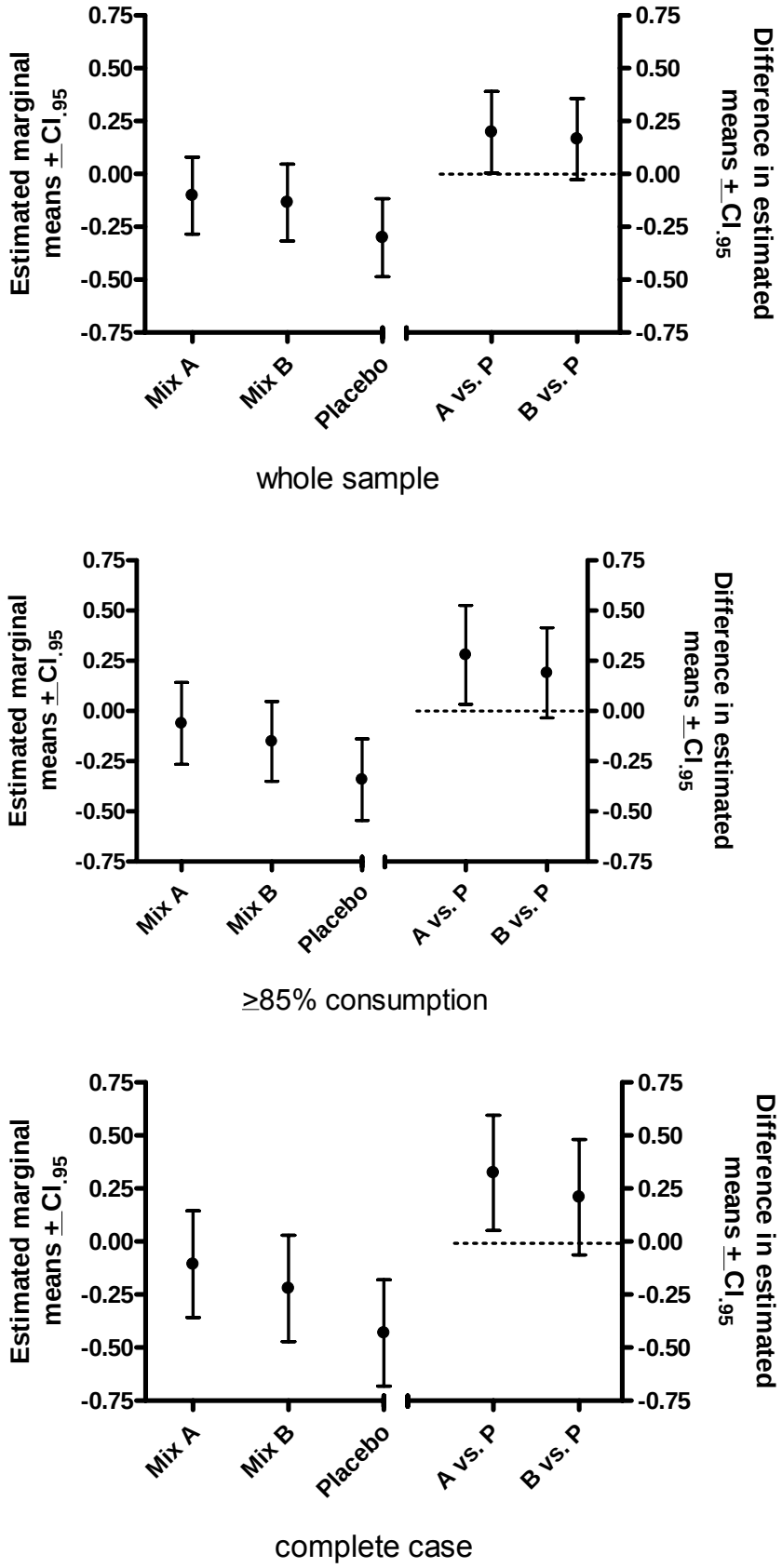


Figure 3 Estimated marginal means by challenge type and the estimated mean difference in GHA on Mix A vs. Placebo and Mix B vs. Placebo under Model 2 for 3 year old children for whole sample (n=153), children with ≥85% consumption in any challenge week (n=132) and complete case group of children with ≥85% consumption over all challenge weeks and no missing GHA data (n=73)

8/9 year olds

Study sample

Of 144 children (mean age 106.3 months, SD: 5.9, range: 93 to 123) enlisted to the study, 75 were boys (mean age 106.4 months, SD: 6.1, range: 93 to 123) and 69 girls (mean age 106.1 months, SD: 5.8, range: 96 to 116). Table 1 provides characteristics separately for the primary outcome, i.e. the whole sample (n=144), a sub-group of children (n=119) with what we defined upon reviewing the results as an acceptable (≥85%) consumption of drinks and a 'complete case' group with ≥85% consumption and no missing GHA data (n=91). No significant differences in these background characteristics were found between sub-samples or between groups of children assigned to receive the challenge drinks in different orders over each of the 6 periods (weeks).

Consumption of drinks

Of those children who dropped out of the study, which occurred for various reasons not linked to problems with the child's behaviour (see Section 3 Annex 2). Two had a mean of 93% consumption in the first challenge week and data was missing for 12 children. Of the remaining children who completed the study, 78% consumed ≥85% of the drinks over the challenge weeks (at least 6 out of 7 daily drinks per week). Only 7 of the remaining 28 children drank less than 50% of placebo and active drinks over the study period.

The occurrence of dietary infractions or 'mistakes' by children over the study period was low (25% = 0; 41% = 1 to 2; 187% = 3 to 4; 16% > 4) (see Annex 2 section 19 for more details). The rate of infractions did not differ during active and placebo weeks.

Effects of challenge

Of 144 children, 125 (86.8%) had full GHA data over active and placebo weeks, 6 (4.2%) had 2 GHA scores, 5 (3.5%) had 1 GHA score and 8 (5.6%) had no GHA scores.

In order to identify potential moderating or confounding effects of a number of variables on the behavioural response of the 8/9 year old children to the challenge mixtures, we examined their relationship to baseline GHA (Table 5). Baseline GHA was significantly related to GHA at all subsequent time points. The significant effects of gender were accounted for by the relationship between gender and baseline GHA, the latter being higher in boys. With baseline GHA in the models, gender was no longer a significant factor. As with the 3 year olds, there was no effect of carryover from challenges in previous weeks on behaviour during subsequent challenge weeks. However there was an effect of week with the GHA increasing with time during the study. This is shown in a trend towards higher GHA with time in these older children. An examination of the components of the GHA indicates that this is due to a gradually worsening of the children's scores on the computerised test of attention. This is an intrinsically boring task on which the children became less motivated with repeated testing.

Table 6 show the unadjusted mean scores of the GHA for the three challenges. Table 7 gives the effect of challenge on GHA (Model 1) and the same analysis repeated but including the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

Table 5: Effect of possible moderators on Baseline GHA for 8/9-year-olds (Total N=

Moderator	Groups Mean (sd)	Mean (sd)	Test
	Age ≤106m (n=57)	Age >106m (n=62)	
Age Gp	.07 (1.02)	-.10 (1.03)	F(1,117)=.749, p=.388
	Male (n=64)	Female (n=55)	
Gender	.35 (1.12)	-.44 (.71)	F(1,117)=20.32, p=.000
	Pre-trial diet (n=119)		
Pre-trial diet	Analysed as a continuous covariate		F(1,117)=.035, p=.852
	O level or below	A level or higher	
Maternal education level	.13 (1.17)	-.15 (.90)	F(1,115)=2.02, p=.158

There were no significant interactions between challenge and these variables. Analyses are shown separately for the full sample (primary outcome of the study), high consumption (case included if ≥85% consumption of drinks in any challenge week) and complete case groups (≥85% consumption in all challenge weeks and no missing GHA) and the estimated marginal means in standard deviation units based on baseline GHA.for the challenges adjusted for the effects of the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2) are displayed in Figure 4.

For the whole sample (N=136) there was a statistically significant effect of Mix B on GHA for both Model 1 (effect size = .12,$CI_{.95}$ .03 to .22, p<.05) and Model 2 (effect size = .12,$CI_{.95}$ .03 to .22, p<.05).

After adjusting for juice consumption (≥85%), the adverse effect of Mix A on behaviour was statistically significant for Model 1 (effect size = .12,$CI_{.95}$ .02 to .23, p<.05) but not for Model 2 (effect size = .09,$CI_{.95}$ -.01 to .19, ns). For this group the effect of Mix B was significant for Model 1 (effect size = .15,$CI_{.95}$ .05 to .25, p<.05) and Model 2 (effect size = .15, $CI_{.95}$ .05 to .25, p<.05) . When the analysis was limited to those with ≥85% juice consumption and no missing GHA data the effects of Mix A were significant for Model 1 (effect size = .14,$CI_{.95}$ .03 to .24, p<.05) and for Model 2 (effect size = .12,$CI_{.95}$ .03 to .23, p<.05) as were the effects of Mix B for Model 1 (effect size = .17,$CI_{.95}$ .06 to .28, p<.01) and Model 2 (effect size = .17,$CI_{.95}$ .07 to .28, p<.01) .

Table 6 Mean GHA for 8/9 year olds by challenge type

	Mix A		Mix B		Placebo	
	n=	**mean (SD)**	**n=**	**mean (SD)**	**n=**	**mean (SD)**
8/9 year olds						
Whole sample n=136	132	0.25 (0.97)	133	0.33 (1.10)	127	0.19 (1.03)
≥85% consumption n=119	104	0.26 (0.93)	112	0.32 (1.09)	103	0.19 (1.04)
Complete case n=91	91	0.27 (0.92)	91	0.35 (1.08)	91	0.19 (1.06)

Table 7 General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	.16 [-.01 to .34]	.09 [-.09 to .27]	.11 [-.10 to .32]
Challenge type			
Mix A –v- placebo	.08 [-.02 to .18]	.12 [.02 to .23] *	.14 [.03 to .24] *
Mix B –v- placebo	.12 [.03 to .22] *	.15 [.05 to .25] **	.17 [.06 to .28] **
Model 2			
Intercept	.02 [-.22 to .26]	.14 [-.08 to .37]	.14 [-.12 to .39]
Challenge type			
Mix A –v- placebo	.08 [-.02 to .17]	.09 [-.01 to .19]	.12 [.02 to .23] *
Mix B –v- placebo	.12 [.03 to .22] *	.15 [.05 to .25] **	.17 [.07 to .28] **
Week of study			
Wk 2 –v- Wk 6	-.11 [. 21 to -.00] *	-.19 [-.29 to -.08] **	-.20 [-.32 to -.09] **
Wk 4 –v- Wk 6	.06 [-.03 to .14]	.04 [-.06 to .13]	.03 [-.07 to .13]
Gender	.16 [-.03 to .35]	.08 [-.10 to .26]	.11 [-.09 to .31]
Baseline GHA score	.78 [.69 to .88] ***	.79 [.71 to .88] ***	.79 [.70 to .89] ***
Pre-trial diet	.04 [-.02 to .10]	.03 [-.03 to .09]	.02 [-.05 to .09]
Maternal education level	-.02 [.20 to .16]	-.02 [-.19 to .15]	.01 [-.18 to .21]
Maternal social class	.04 [-.14 to .22]	-.03 [-.20 to .14]	-.06 [-.25 to .13]

* p<.05, ** p<.01, *** p<.001

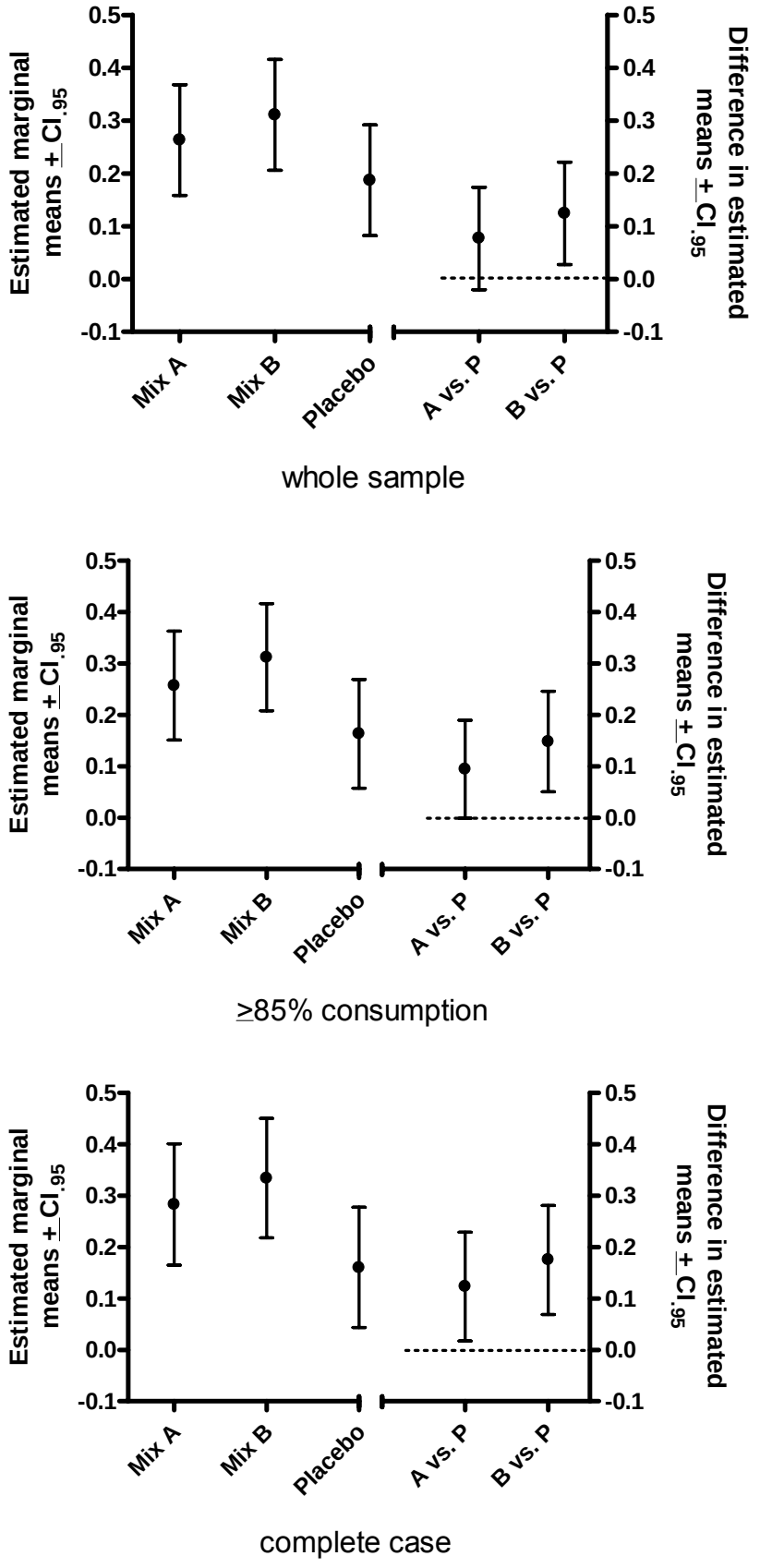


Figure 4. **Estimated marginal means by challenge type and the estimated mean difference in GHA on Mix A vs. Placebo and Mix B vs. Placebo under Model 2 for 8/9 year old children for whole sample (n=136), children with ≥85% consumption in any challenge week (n=119) and complete case group of children with ≥85% consumption over all challenge weeks and no missing GHA data (n=91).**

Discussion of the primary research question finings

The results of this 6 week sub-acute Double Blind Placebo Controlled Food Challenge of the effects of two mixtures of certain artificial food colours and benzoate preservative on children's behaviour, have shown an increase in the mean level of hyperactivity in children aged 3 and 8/9 years. Based on the primary analysis of the data on an intention to treat basis (i.e. the whole cohort), the specific findings were that with Mix A there was, on average, a significant adverse effect on the hyperactive behaviour of 3 year old children as measured using the chosen outcome measure of a Global Hyperactivity Aggregate (GHA). In contrast, Mix B was without significant effect on the behaviour of 3 year old children. The reverse picture was seen with 8/9 year old children. In this case, compared with placebo, Mix B had a significant adverse effect on the behaviour of children. However, Mix A was without significant effect on the behaviour of 8/9 year old children.

Using the complete case data the effect sizes, in terms of the difference between the mean GHA under active Mix (A or B) and placebo challenges, were very similar for Mix B in 3 year olds and the 8 year olds reported here (ES=.21 and .17 respectively). For Mix A the effect for 3 year olds was greater (.32) than for 8 year olds (.12). The reason why the effects for Mix B were not significant for the 3 years olds may be because there was greater variability in the response to the active challenges compared to placebo in these younger children. As this suggests, there are marked individual differences in the response of children to the mixtures, such that not all children responded to the additive mixtures in the same direction, and some children showed no response at all. For these 3 and 8/9 year old children there were no statistically significant effects of baseline GHA, gender, baseline GHA score, pre-trial additive content of diet or social and demographic factors in moderating the effects of challenges.

This study provides, for the first time, evidence of deleterious effects of certain mixtures of artificial food colours and benzoate preservative on children's behaviour with data from a general population sample, using robust objective measures with strong ecological validity, based in part on observations in the classroom and applying double blind challenges with quantities of additives equivalent to realistic dietary intakes. The findings also replicate the effects of Mix A previously reported on a large sample (n=277) of 3 year olds[.].

The specific deleterious compounds in the mix cannot be determined for the present study and need to be examined in subsequent studies. It would require a large and complex study to isolate the effects of individual components of the mixes we applied. Such a study would also have to take into account the possible interactions between the particular additives concerned. These interactions may be particularly important since there is some evidence from in vitro studies, that certain mixtures of other types of additives, namely the combination of the sweetener aspartame and certain colours, is particularly potent in influencing neural development. [23] Significant synergy was observed between combinations of Brilliant Blue (E133) with L-glutamic acid (E620), and Quinoline Yellow (E104) with aspartame (E951) in the inhibition of the growth of neurites in mouse NB2a neuroblastoma cells. Only one of these colours (Quinoline Yellow) was present in the mixes used in the present study (in Mix B).

The present findings in combination with the replicated evidence for the effects of a mixture of certain artificial food colours and benzoate preservative on the behaviour of 3 year old children, provides support for the case that certain food additives may exacerbate

hyperactive behaviours (inattention, impulsivity and overactivity) in some groups of children, at least up to middle childhood. We consider that these findings demonstrate that adverse effects are not just found in those at the extreme of hyperactivity, namely those diagnosed with ADHD,[4] but can also be found in the general population and across the range of severities of hyperactivity. Our results are consistent with those from previous studies and extend the findings to demonstrate significant effects to the general population. The effects are shown after a rigorous control of placebo effects and for children with the full range of levels of hyperactivity.

Possible Implications for public health

Using the results for the children taking 85% or more of the challenge drinks, the effect sizes observed in these general population samples for three year olds were .28 and .19 and for 8-9 year olds .09 and .15 for Mix A and Mix B respectively. These estimates are higher for the three year olds but also only statistically significant for Mix A in 3 year olds and Mix B in 8 to 9 year olds. We consider that the reason for this is that the variability in the differential responses to the mixes compared to placebo was greater in the younger children. In this study the effect sizes average at about 0.18. This is lower than that previously found for clinical groups (0.28).[4]

If the effects of additives hold across the range of levels of hyperactivity, then we hypothesise that removal of these artificial food colours and sodium benzoate preservative with an effect size of 0.18 may lower the population mean. At the extreme the percentage of children scoring more than 1.5 SD above we predict that the mean (6.6%, a typical population prevalence for ADHD[8]) might be lowered to 4.6%. If this were the case, it would result in a 30% reduction in the prevalence of ADHD in children.

Secondary research questions

Effects of the two mixtures of food colour and benzoate preservative on component measures of the GHA

The study was designed with an aggregated measure of hyperactivity as the primary outcome measure and accordingly the results presented in Annexes 1 and 2 are mainly based on quantifying the observed changes in this measure in response to the challenge mixtures compared with placebo. A feature of the earlier work on the Isle of Wight which added uncertainty to the interpretation of the results, was that the significant effects of the artificial food colour and benzoate preservative mixture that was Mix A were only detected by parental reports. The results for the individual component measures of behavioural assessment (i.e. the disaggregated measures), for the whole cohort, for the subset of children consuming 85% or more of the drinks, and for the complete case group of children consuming more than 85% of the drinks, in each age group in the present study are summarised in Table 8. Here the analyses presented for the GHA are replicated for the individual components of the GHA – i.e. testing for the effect of challenge alone (Model 1) and the with covariates (Model 2). Full details of these analyses are in Section 21 of Annexes 1 and 2.

For the whole sample of 3-year-olds the largest effect was that based on parental report for both Mix A and Mix B. By contrast for the whole sample of 8/9-year olds the largest effects were found for the computerised test of attention (CPTII), although for this whole sample analysis only parent report for Mix B was significant. Indeed the significant effect

for GHA under Mix A for 3 –year olds is mainly determined by parental report but the significant effect of Mix B for the 8/9-year-olds is based on both parent rating and the CPTII with the largest contribution coming from the latter.

For the ≥85% consumption and the complete case samples there is a trend for the effect size for parent report to be greater than for whole sample. A similar trend is seen for the CPTII for 8/9 year olds. It should be noted that the effect sizes for computerised test for 8/9 year olds only reach significance for the compete case sample but have values for the whole sample (.19) that are as great as the significant effects for Mix A for 3 year olds for the GHA (.20). It is the greater variability, influenced in part by the greater measurement error that precluded this disaggregated indicator from reaching significance more often.

These results confirm our judgement made, when the study was designed, that hyperactivity is a behaviour that is best measured using assessments from a number of sources. It is a principle in measurement theory that any single indicator is likely to be less valid and relatively unreliable compared to an aggregate measure. For this reason the effect sizes are likely to be lower and the increased measurement error makes it less likely that a significant effect will be detected. The majority of the effect sizes are in the direction of hyperactivity being more marked under the active than the placebo challenge, and are statistically significant for the parental ratings for 3-year-olds and for 8/9 year olds and for the computerised test of attention in the 8/9years. As before, for the younger children, the strongest effects are found for parental ratings. An insight into why this might occur has come from our analysis of an acute challenge.

Table 8 Effect sizes for disaggregated hyperactivity measures over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Effect size (p if p<.10)					
	Mix A vs. Placebo			Mix B vs. Placebo		
	Whole sample	≥ 85% consumption	Complete case	Whole sample	≥ 85% consumption	Complete case
3-year-olds						
Parent rating	.33 (p=.058)	.49 (p=.016)	.55 (p=.027)	.27	.36 (p=.079)	.37
Teacher rating	.01	.03	.09	.06	.08	.10
Classroom Observation	.09	.10	.08	.001	-.01	-.02
8/9-year-olds						
Parent rating	.01	.03	.03	.13 (p=.031)	.13 (p=.046)	.08
Teacher rating	-.04	-.01	.00	-.03	.01	.04
Classroom Observation	.02	.08	.04	.01	.05	.07
Continuous Performance Task	.10	.08	.18	.19	.20	.32 (p=.015)

Acute challenge

The acute challenge was a “proof of principle” component of the project exploring the possibility of demonstrating short term changes in hyperactive behaviour immediately post challenge. We identified two groups of 15 eight year old boys who did or did not respond to Mix B in the main challenge (see Annex 2 Section 22 for details of how these groups were defined). They were brought into the lab on two occasions a week apart and given on each occasion in a random sequence either an active or a placebo challenge. Their

behaviour was monitored using independent observer ratings and the Continuous Performance Task. These were combined in a Hyperactivity Index (HI) and behaviour change calculated between the periods before and after the placebo and active challenges. The results shown in the Figure 5 suggest that responders do show a greater exacerbation of hyperactivity on Mix B compared to placebo than the non-responders although this was not statistically significant (Interaction challenge type x responder status: p<.073). The results from the acute study would be expected to be consistent with the main study. The responders should show a greater increase in hyperactivity under Mix B than placebo and the non-responders would have changes in hyperactivity under Placebo the same as or lower that those for Mix B. This was the pattern that was found.

The findings form the acute challenge study suggest that the 8/9 year old children who respond to the ingestion of Mix B additives do so within a short period of time (an hour). The children in the main study were asked to take their drink challenges at home and this was usually on their return from school. The acute effect that these results suggest therefore makes it more likely that parents in the home setting will be exposed to the behavioural changes and may be an explanation as to why the effects were detected in the parental ratings rather than the school or early years settings based measures.

The full details of the acute challenge study can be found in Section 23 of Annex 2 and details of the analyses in Table 23.1

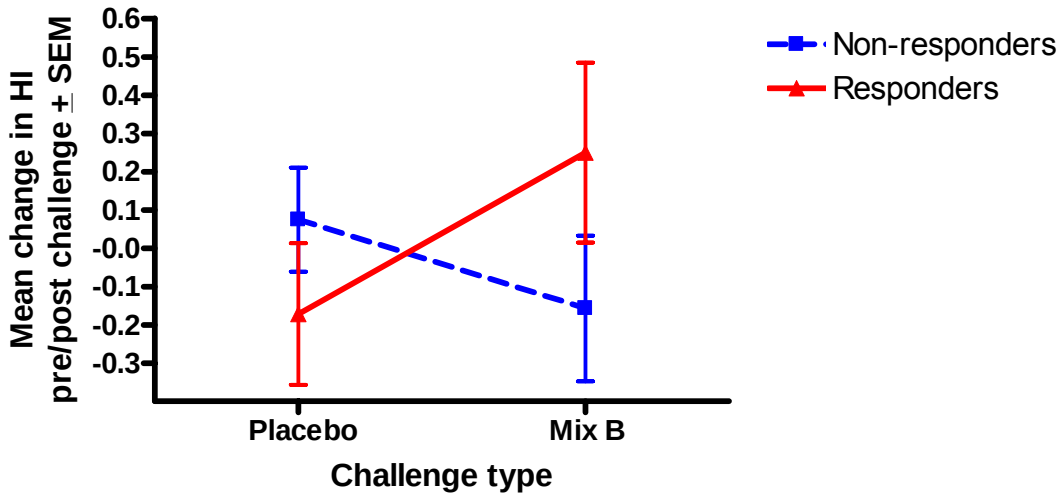


Figure 5. Change in Hyperactivity Index (HI) pre to post challenge in the acute challenge study for eight year old responders and non-responders from main study

Consistency in response between Mix A and Mix B in main study

The mixes differed from one another both in terms of the artificial food colours included and in their doses. It is therefore difficult to interpret differential responses by individual children to the mixes. The metric of the response to a mix was calculated as the difference between the child's Global Hyperactivity Aggregate (GHA) on the mix minus the GHA on placebo. The distribution of these difference scores for both mixes at both ages was normal i.e. there was no immediate evidence from the distribution of a sub-group who were distinctively responsive to the mixtures – the effects were on a continuum. An arbitrary definition therefore had to be used to identify "responders" i.e. being in the top quartile.

Of the 34 three year old children who were responders to Mix A 19 (56%) were responders to Mix B as well. The association between being a responder to Mix A and to Mix B was highly significant (χ^2 = 20.60, d.f. = 1, p <.0001). A similar pattern was seen for eight year olds with 15 (48%) of the 32 Mix A responders being responders to Mix B as well (χ^2 = 9.92, d.f. = 1, p <.002).

Across all children the general pattern of the results we obtained was for both mixes to result in an increase in the mean level of hyperactive behaviour at both ages but with some effects falling short of statistical significance. The effects were more marked (greater effect sizes) for three year olds than for eight year olds, especially for Mix A, although there was significant individual variation in response within each group. The dose, in terms of mg of AFCAs per kg body weight (using average body weights for age), was higher for Mix A in three year olds (1.39mg/kg) than in eight year olds (0.81mg/kg). The mg/kg dose for Mix B was similar for both ages (2.08 and 2.01).

The study design only allows a test for the overall effect of each mixture not for its components. The interpretation of the results needs to recognise the possible existence of interactions between the additives in the mixtures which an in vitro study using different additive combinations to those in this study has indicated to influence neural development.[23] It should be noted that there are marked individual differences between children in the extent to which they respond. This is especially true of the three year olds. As we show below this is in part accounted for by genotypic difference between the children.

Dose-response relationship

We have not undertaken an extensive dose-response analysis as the study was not designed to be able to investigate the effect of dose on response. The measurements of dose in a community based trial such as this are therefore necessarily crude. Perhaps more importantly the variations in dose taken for each mix were not under experimental control but rather arose from differential compliance with the stipulated dose. This means that the interpretation of the dose-response relationship is problematic not least in terms of ambiguity in the direction of effects.

Difference between 3 and 8-9 year olds in the changes in GHA over time

As a result of the design requirements from the FSA, after a baseline period on normal diet, the children were placed on a withdrawal diet without additives and simultaneously started on a placebo drink. This was done to minimise placebo effects so that that through out the study (after baseline) the children were receiving a drink of some kind. This meant that the effects of withdrawal were confounded with those of placebo. The pattern of changes in GHA over the period of the study was that for the three year olds the scores were below the baseline level (i.e. less hyperactive). This we interpret as the effects of withdrawal not being counteracted by the effects of subsequent challenges. For the eight year olds the GHA tended to increase above the baseline level (i.e. more hyperactive), even during periods of placebo challenge. We consider that the difference between the patterns of scores for the two age groups lies in the inclusion of the Continuous Performance Test for the eight year olds. This component of the GHA (which was absent for three year olds) showed progressively worsening scores over the period of the study. The test is intrinsically uninteresting and on repeated assessment the eight year olds became increasingly less motivated and their performance declined week by week (resulting in increased GHA scores). All the other components of the eight year olds' GHA remained below or close to the baseline values. It should be noted that these time effects were controlled in the mixed model analyses we have reported.[6]

Moderation of the effects of food additives by genetic polymorphisms

Hyperactivity in children is characterized by inattention, impulsivity and overactivity. There are marked individual differences in this behavior in the general population. A diagnosis of ADHD is usually reserved for those children with a severe degree of hyperactivity and with a pervasive pattern of behavior from a young age which impairs other aspects of functioning, for example at school.[24]

Variations in hyperactivity have also been attributed to non-genetic effects – with prematurity,[25] institutionalized rearing[26] and maternal smoking during pregnancy[27] all implicated. Increased hyperactivity is an outcome of a wide range of adverse experiential factors that involve central nervous system (CNS) damage.[28] There is, however, a substantial body of evidence that genetic factors are a major contributor to these individual differences in hyperactivity. For example, numerous twin studies indicate that approximately two thirds of the variance in hyperactivity can be explained by genetic differences between children and molecular genetic studies have implicated a number of genes as contributing to this effect.[7]The main group of genes identified to date is that influencing the dopamine system (e.g. DRD4 and DAT1). Other genes in the serotonin and noradrenergic neurotransmitter systems have also been implicated. However, the size of the effects of the genes identified to date are such that they account both individually and in aggregate for only a small portion of the genetic risk suggested by quantitative genetic analyses. It is possible that the effects of these genes will only be apparent in association with environmental and experiential influences. It may be that genes influencing other neurotransmitters systems also contribute.

Histamine is an interesting candidate neurotransmitter system for a number of reasons. The activity of central histamine H-3 receptors have been shown to affect inhibition learning, increase hyperactivity levels in mouse models and to promote dopamine release in frontal cortex[29]. There is evidence that histamine might mediate the effects of AFCAs on hyperactivity. Azo dyes have been shown to provoke urticaria in a minority of individuals with chronic urticaria, independent of whether or not they are aspirin sensitive, providing clinical evidence that artificial colours may result in histamine release. The same study found raised concentrations of urinary and plasma histamine following challenge with tartrazine but not with carmoisine, sunset yellow and amaranth[30]. This is supported by a study of the addition of azo dyes to an *in vitro* system containing circulating basophils which caused a non-IgE dependent histamine release[31]. It has been proposed that the effect of food additives is likely to be a non-specific pharmacological effect that would be similar in children irrespective of their atopic status or other characteristics[32]. Indeed in the previous study on 3 year olds there was no moderating effect of atopy on the elevation of hyperactivity when children were given AFCAs[5]. For these reasons histamine release and its effects on the CNS may play a crucial role in mediating the effects of AFCAs on hyperactivity.

The analysis examined whether the effect of AFCAs on hyperactivity is moderated by genetic difference between children. The results have been prepared for publication.[33].Such an investigation may help to identify which processes mediate the effect of AFCAs on hyperactivity. Genetic polymorphisms were selected from the dopamine (catechol-o-methyltransferase) and adrenergic (adrenergic receptor alpha 2A) neurotransmitter systems since these have previously been implicated in ADHD. Since

there is a suggestion that histamine may be involved in any effects of artificial food colours and benzoate presservative, genetic polymorphisms from this system were also included. Given the difficulties in genotyping variable number tandem repeat (VNTR) polymorphisms using DNA from cheek cells (the only means of access to DNA for general population samples of children) polymorphisms were selected for the present analysis that were single nucleotide polymorphisms (SNPs). Consequently results are presented here for two SNPs in the histamine N-methyltranferase gene, Thr105Ile (rs1801105) and T939C (rs1050891), for one SNP in the dopamine gene, catechol-o-methyltransferase Val108Met (rs4680), and one SNP in the adrenergic neurotransmitter system adrenergic receptor, alpha 2A -1291>G (ADRA2A C1291G, rs1800544).

The following reports the results of analyses of whether there are moderating effects of these genotypes on the effects of the artificial food colour and benzoate preservative mixtures (A and B). The effect of challenge (Mix A vs. Placebo; Mix B vs. Placebo) on GHA was tested using mixed models analyses which used a compound symmetry model for 3 year olds and an unstructured model for the 8/9 year olds. The analysis was limited to those consuming an adequate amount of the challenge since the aim of this analysis was not to establish the impact of the additives per se (where the intention to treat based on the whole sample is the focus). Instead this analysis tries to identify factors modifying this effect. For this purpose the" at least 85% consumption" sub-sample is optimal since close to the full challenge is being consumed and more robust effects should therefore be identified for which moderation can be tested.

Table 9 and 10 present the effects of genotype on GHA at baseline in children consuming $\geq$85% of the challenge drinks for 3 year olds and 8/9-year-olds respectively. There are no significant effects in these baseline analyses suggesting no main effects of genotype on GHA. During the challenge study however HNMThr105Ile and HNMT T939C were related to the general level of the GHA for 3 year-olds but not for the 8/9 year olds. (see Tables 15.2. and 15.3 in Section 15 Annexes 1 and 2). There were no main effects of COMT Val108Met or ADRA2A C1291G on GHA at either age. For this study the interest did not lie in these main effects but in the interactions between genotype and effects of Mix A and Mix B vs Placebo.

These results are presented in Figure 7 and 8 and are based on examining the moderation of the effect of challenge by the child's genotype for the $\geq$ 85% consumption group with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2). The full details of these analyses are given in Section 18 of Annexes 1 and 2.

The results show that there is no moderation of the effects of additives by the COMT Val108Met and ADRA2A C1291G polymorphism. However such moderation was found for the HNMT Thr105Ile and HNMT T939C polymorphisms. For the three year old children a moderating effect of the 105Ile present genotype was found which significantly reduced the adverse effect of Mix A (p=.041). A similar effect with three year olds for 939C present genotype fell just short of significance (p=.061). The same moderating effect for 105Ile present and 939C present were found for 8/9-year-olds but with Mix B (p=.048 and .026 respectively). In addition 939C present significantly reduced the effect of Mix A (p=.021)

For 3 and for 8/9-year-olds there were no significant moderating effects of COMT Val108Met or ADRA2A C1291G.

Table 9: Baseline analyses of Genotype 3-year-olds

Moderator	n	All Mean (sd)	Groups Mean (sd)	Mean (sd)	Test
HNMT Thr 105Ile	128	-.05 (.68)	105Ile present (n=36) -.11 (.65)	105Ile absent (n=92) -.12 (.68)	F(1,126)=2.88, p=.092
HNMT T939C	125	-.04 (.68)	939C present (n=47) .10 (.60)	939C absent (n=78) -.13 (.71)	F(1,123)=3.41, p=.067
COMT Val108Met	127	-.05 (.67)	Val108 present (n=95) -.06 (.68)	Val108 absent (n=32) .002 (.65)	F(1,125)=.21, p=.646
ADRA2A C1291G	126	-.06 (.68)	1291G present (n=66) -.13 (.68)	1291G absent (n=60) .02 (.68)	F(1,124)=1.56, p=.213

Table 10 Effect of Genotype on Baseline GHA 8/9-year-olds

Moderator	N	All Mean (sd)	Groups Mean (sd)	Mean (sd)	Test
HNMT Thr105Ile	118	-.02 (1.03)	105Ile present (n=24) -.25 (.75)	105Ile absent (n=94) .04 (1.08)	F(1,116)=1.54, p=.217
HNMT T939C	118	-.01 (1.03)	939C present (n=49) .11 (1.19)	939C absent (n=69) -.09 (.89)	F(1,116)=1.03, p=.311
COMT Val108Met	119	-.02 (1.03)	Val108 present (n=85) -.03 (.92)	Val108 absent (n=34) .01 (1.27)	F(1,117)=.046, p=.830
ADRA2A C1291G	116	-.001 (1.03)	1291G present (n=51) -.08 (1.08)	1291G absent (n=65) .06 (1.00)	F(1,114)=.514, p=.475

Fig. 7 The mean GHA by challenge and genotype is shown for 3 year olds taking ≥ 85% of challenge drinks with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

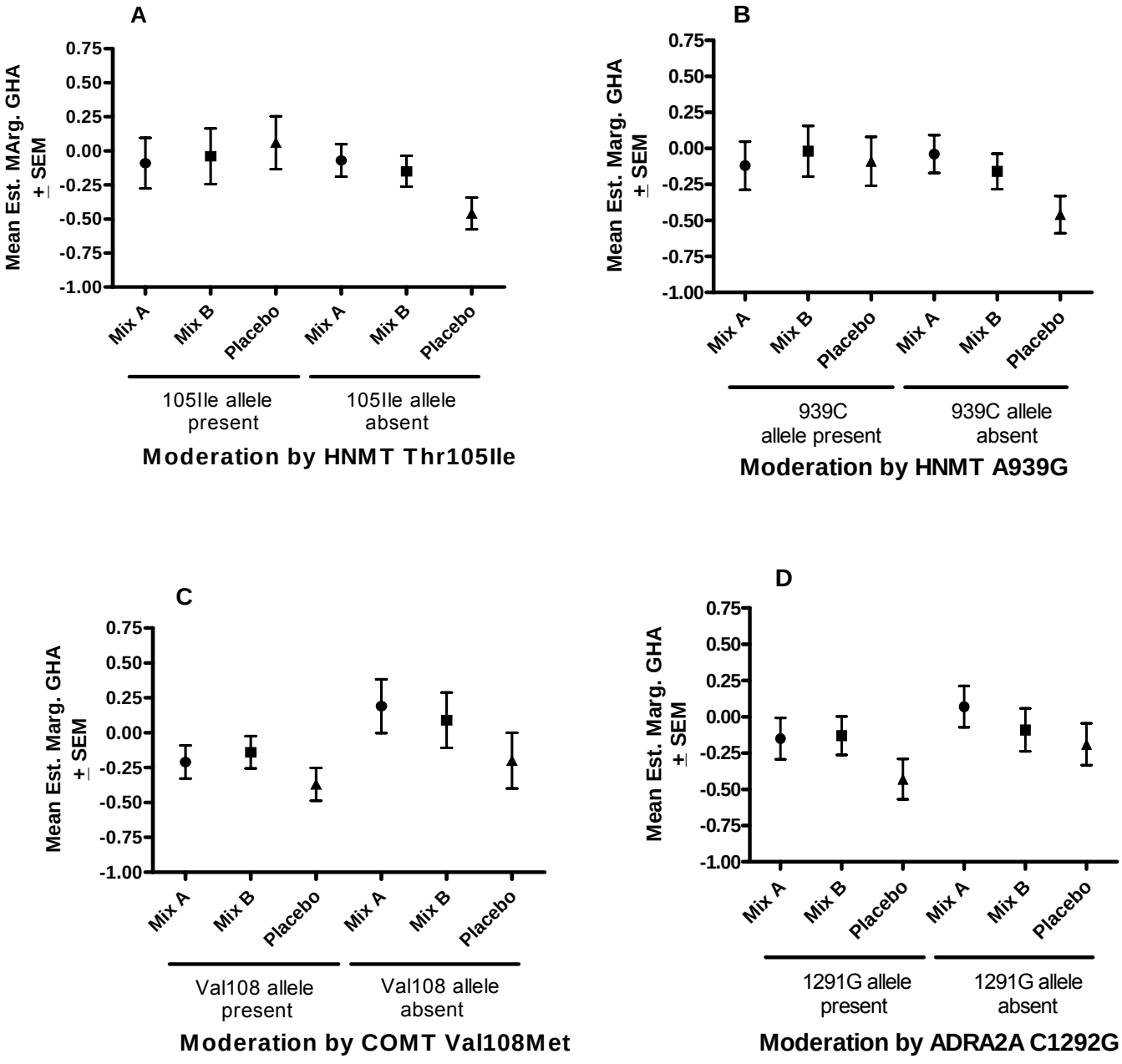
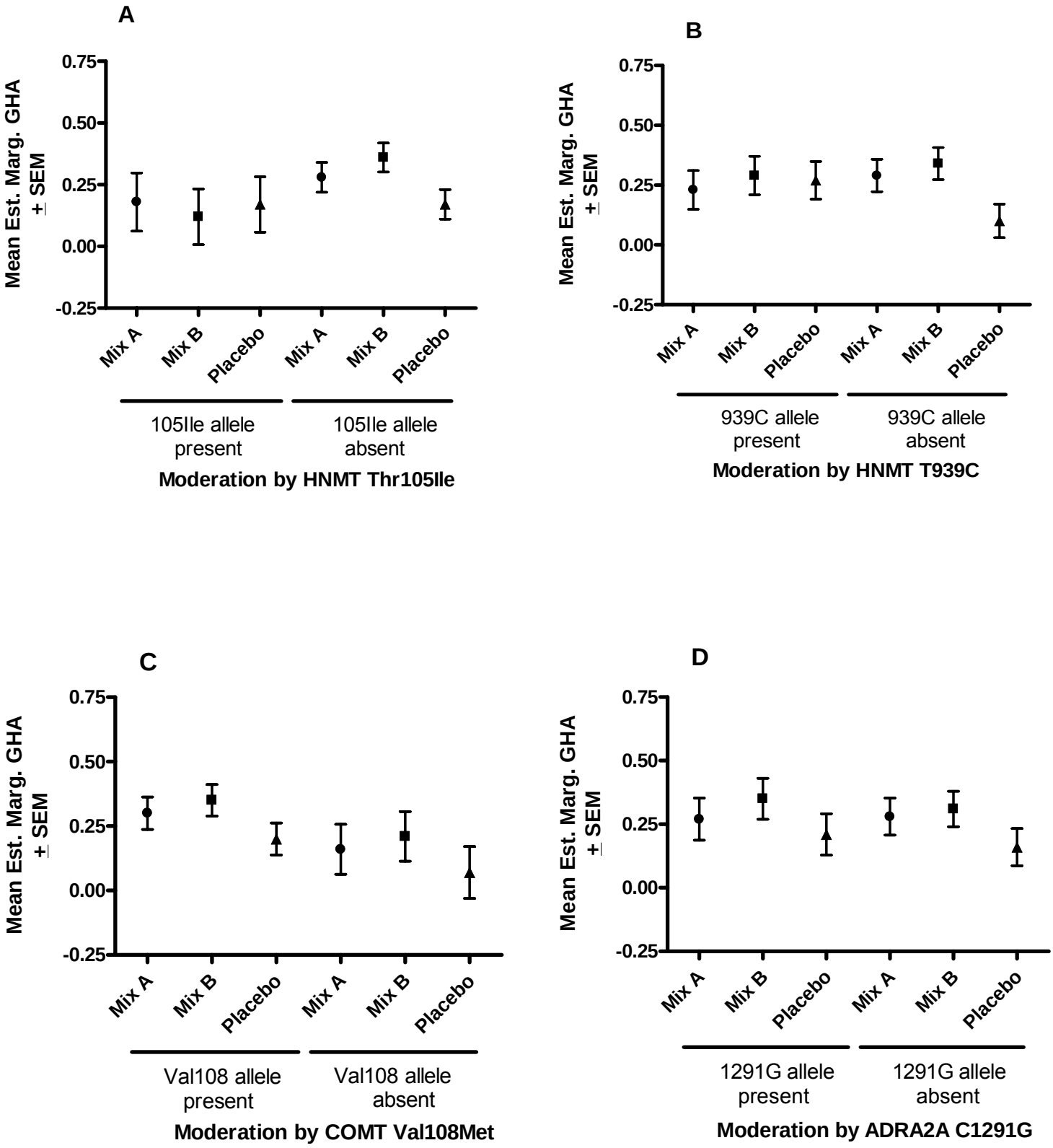


Fig. 8 the mean GHA by challenge and genotype is shown for 8/9 year olds taking ≥ 85% of challenge drinks with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2)



The findings from this study suggest a link between histamine and hyperactivity with certain polymorphisms in the HNMT gene moderating behavioural responses to the mixture of colours and benozoate preservative present in Mix A and B, in the older age group and for Mix A only in the 3-year-olds. The degree of moderation is substantial. For those with the 105Ile present phenotype and the 939C present genotype the effects of the additives seem to be eradicated. For these groups the lines of means in Figures 7 and 8 are essentially flat.

This suggests that the current focus on dopamine in studies of ADHD needs to be extended to histamine. These findings open explanations as to why the genes so far shown to be associated with ADHD explain so little of the known variance. The histamine risk alleles in this study have two actions. The first is to influence the overall level of the GHA in the study, significantly so for the younger children, and second to make the child more vulnerable to the effects of AFCAs in the diet on behavior. The role of genes in influencing behavior needs to be understood not by just their main effects of raising levels, for example, of hyperactivity but also by the interplay[34] both with each other in gene-gene interactions and also by interactions with environmental factors such as diet.

HNMT polymorphisms impair histamine clearance[35] and it is know that challenge with certain artificial food colours can induce histamine release[31] and therefore the interaction we found would be expected. The presence of H3 receptors in the brain provides a potential mechanistic explanation for the effect.[36] Many environmental factors will increase histamine release, e.g. infections as will many food items. This would explain the frequent claim that food allergy/intolerance is a cause of hyperactivity and the effects of infections in aggravating aberrant behaviour. This clearly indicates a potential target for therapeutic intervention in ADHD focused on the H3 receptor.

Conclusions - overall

This study has provided evidence that adverse effects of certain mixtures of artificial food colours and benzoate preservative on hyperactivity can be identified in community samples of 3 year and 8-9 year old children under some circumstances. The size of the effects of additives on the average hyperactivity score is lower than that reported for clinical samples and the level of individual variation in response was high. We recognise that hyperactivity is a behaviour influenced by a wide range of experiential and biological factors. It is known that there are major genetic influences on hyperactivity and the study has shown that differential sensitivity to additives resulting from genetic polymorphisms is one route by which genetic influences on hyperactivity may be mediated. Although the results of the study suggest that some mixtures of certain artificial food colours and benzoate preservative may affect the level of hyperactive behaviour in some children, removal of these additives is not a panacea for ADHD. However, as one risk factor amongst a wide range of risk factors the removal of additives from the diet for these children is likely to be of benefit.

The question of whether the additives in these mixes should be removed from children's food is less clear. This is a question that cannot be decided on the basis of this study alone. What we consider that this double-blind placebo cross-over food challenge has provided, is evidence based on an experiment that the link between additives and hyperactivity is causal – the evidence is not just that of a correlation in survey. It is

possible such an action would have a tangible benefit for public health in decreasing the level of hyperactive behaviours in the community at large and reducing the number of children with dysfunctional levels of hyperactivity.

We have found an adverse effect of some mixtures of artificial food colours and a preservative food additives on the hyperactive behaviour of some groups of 3 and 8/9 year old children. These results replicate the findings published form our previous study.[5] While the use of artificial food colouring might seem superfluous, the same cannot be said for sodium benzoate which has an important preservative function. We consider that the implications of the results for the regulation of food additive use may be considerable.

g) Acknowledgements

We gratefully acknowledge the help and assistance received from the children, families and teachers in the participating EYS and Schools in the Southampton area. We are grateful to the local Steering Committee for assistance with the study especially Chris Talbot, Karen Morris and Ulrike Munford. The study benefited considerably for the advice received from Ian Kimber, Sue Hattersley and her team on behalf of the Food Standards Agency. We are grateful for the funding support of the Food Standards Agency (Grant Ref. T07040).

The full list of names of the FABiC research team is follows. As PI, Prof. Stevenson would like to acknowledge the magnificent effort put in by the team to complete on time a complex and demanding controlled trial of this kind in two large community based samples:

Prof. Jim Stevenson (PI)
Prof. John Warner (Co-investigator))
Prof. Edmund Sonuga-Barke (Co-investigator)
Dr. Donna McCann (Senior Research Fellow)
Kate Grimshaw (Senior Dietitian)
Yuet-Wan Lok (Dietitian)
Debbie Crumpler (Study administrator)
Angelina Barrett (Research assistant)
Alison Cooper (Research assistant)
Lindy Dalen (Research assistant)
Elizabeth Kitchin (Research assistant)
Lucy Porteous (Research assistant)
Emily Prince (Research assistant)
Jenny Scoles (Research assistant)
Catherine Varcoe-Baylis (Research assistant)

h) References

1 Feingold BF. Hyperkinesis and learning disabilities linked to artificial food flavors and colours. Am J Nurs 1975; **75:** 797–803.

2 Overmeyer S, Taylor E, Annotation: principles of treatment for hyperkinetic disorder: practice approaches for the UK. J Child Psychol Psychiatry 1999; **40:** 1147–57.

3 Editorial. NIH consensus development conference: defined diets and childhood hyperactivity. Clinical Paediatrics 1982; **21:** 627–30.

4 Schab DW, Trinh NT. Do artificial food colours promote hyperactivity in children with hyperactive syndromes? A meta-analysis of double-blind placebo-controlled trials. J Dev Behav Pediatr 2004; **25:** 423–34.

5 Bateman B, Warner JO, Hutchinson E, et al. The effects of a double blind, placebo controlled, artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. Arch Dis Child 2004; **89:** 506–11.

6 McCann D, Barrett A., Cooper A, Crumpler D, Dalen L, Grimshaw K, Kitchen E, Lok K, Porteous L, Prince E, Sonuga-Barke E, Warner JO, Stevenson J. Food additives and hyperactive behaviour in 3 and 8/9 year old children in the community: a randomised, double-blinded, placebo controlled trial.. The Lancet, September 6 2007, DOI:10.1016/S0140-6736(07)61306-3

7. Waldman ID, Gizer IR . The genetics of attention deficit hyperactivity disorder**.** Clini Psychol Revi 2006: **26**: 396-432.

8 Swanson J.M., Sergeant J., Taylor, E., Sonuga-Barke E., Jensen P.S., Cantwell D. Attention Deficit Hyperactivity Disorder and Hyperkinetic Disorder. The Lancet, 1998;**351**: 429-433.

9 DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey K, Ikeda M. Teacher ratings of ADHD symptoms: Factor structure and normative data. Psychol Assess 1997; **9:** 436–44.

10 Gregory JR, Collins EDI Davies PSW, Hughes JM, Clarke PC. National Diet and Nutrition Survey: children aged 1·5 to 4·5 years. Vol 1: Report of the Diet and Nutrition Survey. London: HM Stationery Office, 1995.

11 Egger J, Graham PJ, Carter CM, Gumley D, Soothill JF. Controlled trial of oligoantigenic treatment in the hyperkinetic syndrome. Lancet 1985; **325:** 540–45.

12 Carter CM, Urbanowicz M, Hemsley R, et al. Effects of a few food diet in attention-deficit disorder. Arch Dis Child 1993; **69:** 564–68.

13 Routh D. Hyperactivity. In: Magrab P, ed. Psychological management of pediatric problems. Baltimore: University Park Press, 1978: 3–8.

14 Thompson MJJ, Stevenson J, Sonuga-Barke E, et al. The mental health of preschool children and their mothers in a mixed urban/rural population. Prevalence and ecological factors. Br J Psychiatry 1996; **168:** 16–20.

15 Hayward C, Killen J, Kraemer H, et al. Linking self-reported childhood behavioural inhibition to adolescent social phobia. J Am Acad Child Adolesc Psychiatry 1998; **37:** 1308–16.

16 Mash EJ, Johnston C. Parental perceptions of child behaviour problems, parenting self-esteem and mother's reported stress in younger and older hyperactive and normal children. J Consult Clin Psychol 1983; **51:** 86–99.

17 DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. AD/HD rating scale IV: checklists, norms and clinical interpretation. New York: Guilford Press, 1998.

18 Abikoff H, Gittleman R. Classroom observation code—a modification of the stony-brook code. Psychopharmacol Bull 1985; **21:** 901-09.

19 Conners CK. The Conners continuous performance test. Toronto, ON, Canada: Multi-Health Systems, 1994.
20 Epstein N, Erkanli A, Conners CK, Klaric J, Costello JE, Angold A. Relations between continuous performance test performance measures and ADHD behaviours. J Abnorm Child Psychol 2003; **31:** 543–54.
21 Gueorguieva R, Krystal JH. Move over ANOVA: Progress in analysing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. Arch Gen Psychiatry 2004; **61:** 310–17.
22 Mallinckrodt CH, Watkin JG, Molenburghs G, et al. Choice of the primary analysis in longitudinal clinical trials. Pharm Stat 2004; **3:** 161–69.
23. Lau K, McLean WG, Williams DP, Howard CV . (2006). Synergistic interactions between commonly used food additives in a developmental neurotoxicity test. Toxicological Sciences, 90, 178–187.
24 McGee R, Prior M, Williams S, Smart D, Sanson A. The long-term significance of teacher-rated hyperactivity and reading ability in childhood: findings from two longitudinal studies. J Child Psychol Psychiatry 2002; **43:** 1004–17.
25. Bhutta AT, Cleves MA, Casey Ph, Cradock MM, Anand KJS. Cognitive and behavioral outcomes of school-aged children who were born preterm - A meta-analysis. JAMA 2002; **288**,: 728-737.
26. Kreppner JM, O'Connnor TG, Rutter M. Can inattentive/overtatcivty be an institutional deprevation syndrome? J Ab Child Psychol 2001; **29**: 513-528.
27. Schmitz M, Denardin D, Sliva TL, PIanca T, HUtz MH, Faraone S, Rohde LA**.** Smoking during pregnancy and attention-deficit/hyperactivity disorder, predominantly inattentive type: A case-control study. J Am Acad Child Adol Psychiat 2006; **45**:1338-1345.
28. Taylor E. Developmental neuropsychopathology of attention deficit and impulsiveness. Dev Psychopath, 1999; **11**: 607-628.
29. Fox GB, Esbenshade TA, Pan JB, Radek RJ, Krueger KM, Yao BB, Browman KE, Buckley MJ, Ballard ME, Komater VA, Miner H, Zhang M, Fagih R, Rueter LE, Bitner RS, Drescher KU, Wetter J, Marsh K, Lemaire M, Porsolt RD, Bennani YL. Pharmacological properties of ABT-239[4-(2-{2-[(2R)-2-methylpyrrolidinyl]ethyl}-benzofuran-5-yl)benzonitrile]: II. Neurophysiological characterization and broad preclinical efficacy in cognition and schizophrenia of a potent and selective histamine H-3 receptor antagonist. J Pharmacol Exper Therap 2005; **313**: 176-190.
30. Murdoch RD, Pollock I, Young E, Lessof MH. Food additive induced urticaria: studies of mediator release during provocation tests. J Roy Coll Phys 1987; **4**: 262-6.
31. Murdoch RD, Lessof MH, Pollock I, Young E. The effects of food additives on leukocyte histamine release in normal and urticarial subjects. J Roy Coll Phys 1987, **4**:251-6.
32. Pollock I, Warner JO. Effect of artificial food colors on childhood behaviour. Arch Disease Childh 1990; **65**:74-77.
33. Stevenson J, McCann D, Grimshaw K, Holloway J, Sonuga-Barke E, Warner JO. Genotype moderates the impact of food additives on hyperactive behavior in children, Arch Gen Psychiat (awaiting submission for publication).
34, M. Rutter, T.E. Moffitt, A. Caspi, Gene-environment interplay and psychopathology: multiple varieties but real effects . J Child Psychol Psychiat 2006; **47**, 226 -261.
35. Preuss, C.V., Wood, T.C., Szumlanski, C.L, Rfatogianis, R.B., Otterness, D.M., Giarard, B., Scott, M.C., Weinshilboum, R.M. Human histamine N-methyltransferase pharmacogenetics: Common genetic polymorphisms that alter activity . Mol Pharmacol, 1998;**53**: 708-717.
36. Sakurai,E., Sakurai, E., Orelandf, L., Nishiyama, S. Kato, M., Watanabe, T Yanai, K. Evidence for the presence of histamine uptake into the synaptosomes of rat brain Pharmacol 2006; **78**, 72-80.

Annex 1

Food and Behaviour in Children Study

3yo Report to FSA

This set of material describes in detail the design, methods, data analysis and findings of the study of 3 year olds. It is accompanied by a parallel report (Annex 2) for the study of 8/9 year old children.

Annex 1 CONTENTS

1. Foreword

The first part of this report presents data on the enlistment of early years settings (Table 2.1), parents and children (Figure 3.1). Early years settings were approached on the basis of their proximity to schools within the City of Southampton enlisted to the 8-9yo study. Based on data relating to free school meal uptake, these schools were representative of all schools across the Southampton area. All but one of the EYS enlisted were located either in the enlisted schools or were within 1 mile distance of such schools.

Information relating to characteristics of the whole sample and stratified by gender are presented in Table 3.1. However, in Table 3.2 we also present characteristics for a subgroup of children with ≥85% juice consumption over any challenge week and separately for a similar but smaller group (complete case) who had ≥85% recorded consumption of juice over challenge weeks and, in addition, no missing behaviour data. These subgroups have been employed in the Mixed Model analyses more detailed comment of which can be found from Section 10 onwards. This first part of the report, therefore, also provides data on the percentage consumption of juices over the period of the study (Section 5). Summary tables are also provided relating to the child's diet e.g. dietary intake prior to the start of the study (Section 6). This was based on maternal recall of the child's diet in the previous 24 hours and was classified on the basis of the number of foods consumed containing the food additives under study. This factor was taken into account in analyses. Further tables provide data relating to the incidence of dietary infractions ('Whoops') reported by parent. This data was employed in some secondary analyses, one of which focused only on those children with ≥85% consumption of juice and no recorded 'Whoops' (Section 19).

Information is provided on the measures of behaviour employed in this 3yo study (Section 7) together with summary tables relating to the distribution of behaviour data collected prior to, at baseline and over the period of the study (Section 9.3). Behaviour screen data collected from practitioners in early years settings prior to the start of the study indicates that the behaviour ratings for our sample of 3yo children are representative of those for children of the same age in participating early years settings schools (see Section 9).

Comment has been made on the use of Mixed Model methods in data analysis and a discussion of the use of an appropriate covariance matrix structure in the Mixed Models approach can be found in Section 10 prior to discussion and presentation of the findings of

the main Mixed Model analyses and analyses based on Genotype (Section 11 onwards). The additional secondary analyses of interest focussing on 'Whoops' data, disaggregated behaviour measures and responders and non-responders are presented from Section 19 onwards. A list of references can be found at the end of the Annex.

2. Early Years Settings (EYS)

EYS were approached and enlisted on the basis of their proximity to Primary and Junior schools enlisted in the first phase of the study. In this first phase information on free school meal percentage uptake (fsm%) was obtained for all schools in the Southampton City Council area. Schools were then ranked into five groups on the basis of percentage uptake (Group 1: 0-10%, Group 2: 11-20%; Group 3: 21-30%; Group 4: 31-40%; Group 5: 41 to 50%). A proportion of schools within each group were approached and enlisted In order to ensure a study sample reflecting the full range of socio-economic background of children in the area. Table 2 (Column 1) lists the participating schools within each group. In order to enlist pre-schools and children of the same broad socio-demographic spread, EYS with a roll of at least 30 children of the target age and located within aprroximately a one mile area of enlisted schools were approached and invited to participate in the study. Of 122 EYS in the Southampton City area, 30 were approached and 26 agreed to participate. Table 2.1 lists the participating EYS within each school fsm% group. After obtaining parental consent, children of the target ages were screened using behaviour rating scales. Table 2.2 shows dates when the Research Team were present in individual EYS to carry out the study.

Table 2.1 Characteristics of participating EYS

Geographiclly representative schools enlisted based on free school meal uptake (fsm%) 2004	EYS enlisted (n=27) within 1 mile area of school(s) in same fsm Group EYS (distance in miles if > 1 mile)
Group 1 (fsm 0-10%) St. Monica Junior School	Spring Road Pre-school, Weston Church Centre Pre-school, Kanes Hill Pre-school (1.5 miles), St Christopher's Playgroup
Group 2 (fsm 11-20%) Bitterne Park Junior School	Riverside Pre-school, Squirrels Corner
Spring Hill Catholic Primary School	Freemantle (CE) Infant School, Paintpots Nursery and Pre-school
Ludlow Junior School	Woolston Community Pre-school, Bitterne Community Pre-school I, Bitterne Community Pre-school II
Group 3 (fsm 21-30%) Fairisle Junior School	Fairisle Infant and Nursery School, Holy Family Pre-school
Tanners Brook Junior School	Rainbow Pre-school, Regents Park Pre-school, Stephens Early Years Centre, Foundry Lane Community Playgroup
Group 4 (fsm 31-40%) St Denys Primary School	University Day Nursery, Asquith Nursery, Brook Pre-school

Group 5 (fsm 41-50%)	
St Mary's C of E (VC) Primary School	St Mary's (CE) Primary School and Nursery, St John's Infant and Nursery School, Startpoint Northam, Maytree Infant School and Nursery
Newlands Primary School	Sticky Fingers Pre-school, Redbridge Community Pre-school

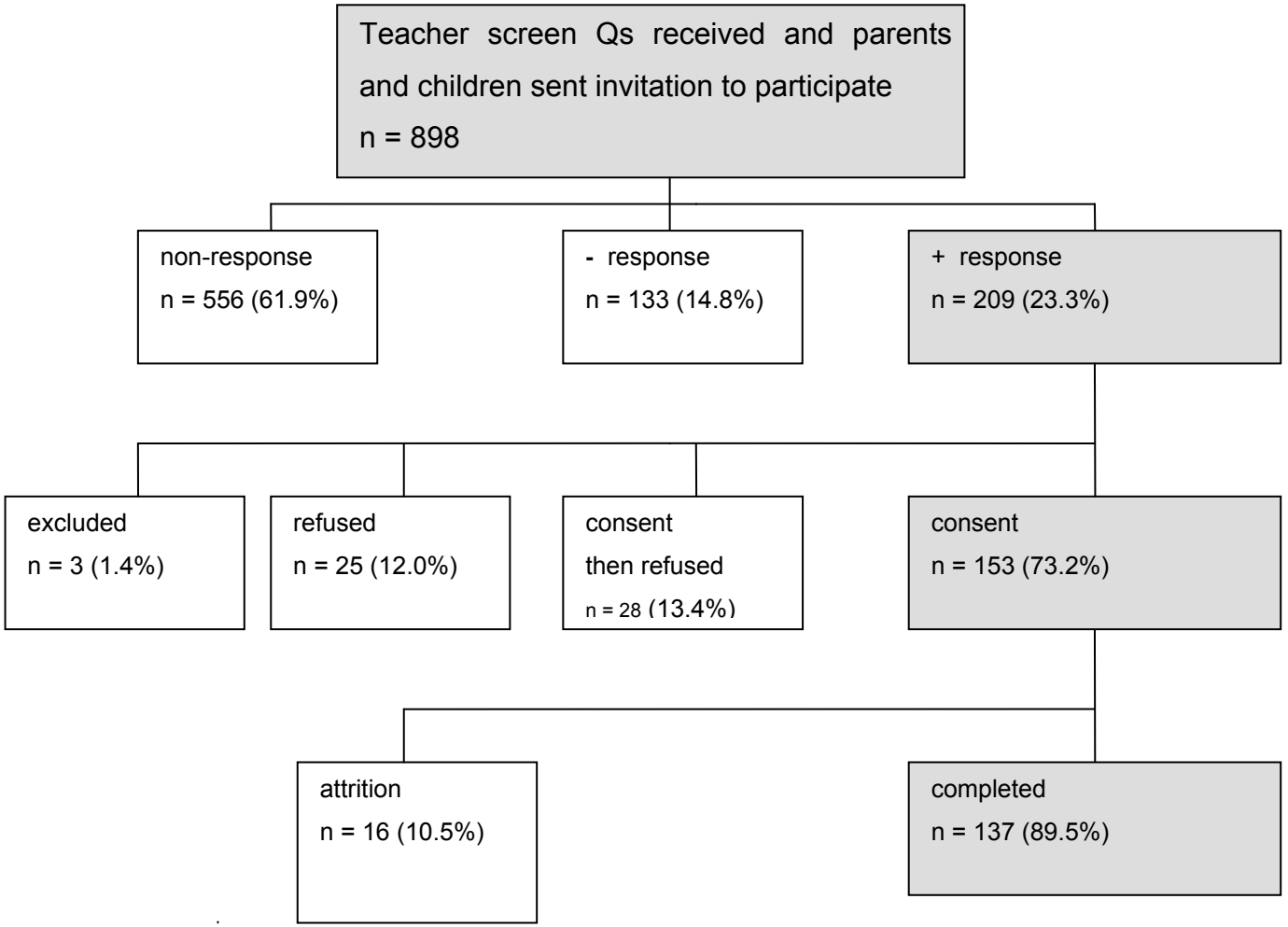
Table 2.2 Periods in EYS

Period 1: Mon 31st October to Fri 16th December 2005: Fairisle Infant and Nursery School (Gp3 – fsm 28%); Regents Park Pre-school (Gp3 – fsm 22%); Riverside Pre-school (Gp2 – fsm 13%); Woolston Community Pre-school I and II (Gp2 – fsm 12%); Sticky Fingers Pre-school (Gp5 – fsm 49%); St Mary's (CE) Primary School and Nursery (Gp5 – fsm 46%); Spring Road Pre-school (Gp1 – fsm 6%)
Period 2: Mon 6th February to Fri 31st March 2006: Woolston Community Pre-school I and II (Gp2 – fsm 12%); University Day Nursery (Gp4 – fsm 31%); Kanes Hill Pre-school (Gp1 – fsm 6%); Rainbow Pre-school (Gp3 – fsm 22%); Brook Pre-school (Gp4 – fsm 31%); Redbridge Community Pre-school (Gp5 – fsm 49%); Bitterne Community Pre School I and II (Gp2 – fsm 12%); Freemantle (CE) Infant School (Gp2 – fsm 11%); St John's Infant and Nursery School (Gp5 – fsm 46%); Foundry Lane Community Playgroup (Gp3 – fsm 22%); Startpoint Northam (Gp5 – fsm 46%)
Period 3: Mon 22nd May to Fri 14th July 2006: Freemantle (CE) Infant School (Gp2 – fsm 11%); Maytree Infant and Nursery School (Gp5 – fsm 46%); Holy Family Pre-school (Gp3 – fsm 28%); Paintpots Nursery and Pre-school (Gp2 – fsm 11%); Asquith Nursery (Gp4 – fsm 31%); Squirrels Corner (Gp2 – fsm 13%); Stephens Early Years Centre (Gp3 – fsm 22%); St Christopher's Playgroup (Gp1 – fsm 6%); Weston Church Centre Pre-school (Gp1 – fsm 6%)

3. Participants

Figure 3.1 provides details of 153 children and their parents enlisted from 26 participating EYS. Of these 137 (89.5%) children completed the study.

Figure 3.1: Enlistment of parents and 3yo children to main study



Behaviour screen questionnaires were completed for children of the target age by practitioners in participating EYS. Parents were subsequently approached via the EYS to participate in the study. Those who returned an expression of interest form were contacted by phone and a convenient time for a home visit arranged. On this visit, a research assistant and the study dietitian provided full information about the study and its dietary implications and informed consent was obtained. At the home visit the study dietitian also obtained a report based on 24-hour recall by the parent of the child's pre-study diet which allowed an assessment of baseline levels of the number of foods containing additives consumed by the child in the previous 24 hour period. Children were invited to sample the juice which would be administered throughout the study.

At this stage, 3 children were excluded from the study as they did not meet study criteria in terms of age (n=2) or were not registered for at least 2 sessions per week at the EYS (n=1 child). A further 25 children were not enlisted to the study or refused to participate prior to or at the home visit (no reason given [n=19], moving home [n=1], no contact obtained [n=4], pressure of work/other commitments [n=1]). Of the 181 parents who consented to participate at the home visit, 28 (15.5%) subsequently refused for reasons related to juice consumption (n=16), medical reasons (n=5), pressure of work/other commitments (n=3), no further contact obtained (n=2), moving house (n=1) or no reason given (n=1). The pre-school children successfully enlisted (n=153) had a mean age of 43.5m (SD 4.5m, range 35 to 53m) with a similar age distribution for the 74 girls (mean age 43.4m, SD 4.4m, range 35 to 53m) and 79 boys (mean 43.5m, SD 4.6m, range 35 to 52m). Table 3.1 presents detail relating to characteristics of all participating parents. For data analysis purposes, Table 3.2 also presents the same information for a subgroup of children (n=133) who had consumed $\geq$85% juice in any challenge week and for a smaller group of these children (n=73) who also had complete aggregated behaviour data. Data analyses are subsequently carried out for all 153 children (intention to treat basis), for those with acceptable levels of juice consumption (n=133) and for a complete case group (n=73). Of the 153 participating children, 16 (10.5%) failed to complete the study for reasons related to juice consumption (n=11), parental pressure of work or other commitments (n=1), no reason given (n=2), medical reasons (n=1) or behaviour related to child (n=1). No differences were found in terms of age, gender or marital status of parent(s) between those children who completed the study and the group of children who did not complete (9 boys, 7 girls).

Table 3.1: Characteristics of parents

	Total n =153 N(%)	Boys N = 79 N(%)	Girls n = 74 n(%)	Diffs $\chi 2$(df)p=
Racial background of parents				
White	126 (82.4)	65 (82.3)	61 (82.4)	
Other	15 (9.8)	7 (8.9)	8 (10.8)	ns
Missing data	12 (7.8)	7 (8.9)	5 (6.8)	
Marital status				
Married/partner	127 (83.0)	66 (83.5)	61 (82.4)	ns
Single/separated/divorced/widowed	26 (17.0)	13 (16.5)	13 (17.6)	
National Statistics Social Class (Father)				
Higher occupations	34 (22.2)	18 (22.8)	16 (21.6)	
Intermediate occupations	26 (17.0)	15 (19.0)	11 (14.9)	
Lower occupations	51 (33.3)	24 (30.4)	27 (36.5)	ns
Never worked/long term unemployed	4 (2.6)	2 (2.5)	2 (2.7)	
No father present	26 (17.0)	13 (16.5)	13 (17.6)	
Missing data	12 (7.8)	7 (8.9)	5 (6.8)	
National Statistics Social Class (Mother)				
Higher occupations	31 (20.3)	16 (20.3)	15 (20.3)	
Intermediate occupations	18 (11.8)	9 (11.4)	9 (12.2)	
Lower occupations	66 (43.1)	34 (43.0)	32 (43.2)	ns

Never worked/long term unemployed	26 (17.0)	13 (16.5)	13 (17.6)	
Missing data	12 (7.8)	7 (8.9)	5 (6.8)	
Mother's educational qualifications				
School to 16 (no qualifications / certificates below 'A' level)	53 (34.6)	27 (34.2)	26 (35.1)	
'A' levels	61 (39.9)	30 (38.0)	31 (41.9)	ns
University Degree/Post-graduate qualification	27 (17.7)	15 (19.0)	12 (16.2)	
Missing data	12 (7.8)	7 (8.9)	5 (6.8)	

Table 3.2: Characteristics of parents of children (all children, ≥ 85% consumption, complete case)

	Children in total sample Analysis n =153 n(%)	Children in ≥85% drunk Analysis n = 133 N(%)	Children in complete case analysis n = 73 n(%)
Racial background			
White	126 (82.4)	117 (88.0)	67 (91.8)
Other	15 (9.8)	12 (9.0)	6 (8.2)
Missing data	12 (7.8)	4 (3.0)	-
Marital status			
Married/partner	127 (83.0)	110 (82.7)	60 (82.2)
Single/separated/divorced/widowed	26 (17.0)	23 (17.3)	13 (17.8)
NSSC* (Father)			
Higher occupations	34 (22.2)	30 (22.6)	15 (20.5)
Intermediate occupations	26 (17.0)	24 (18.0)	9 (12.3)
Lower occupations	51 (33.3)	48 (36.1)	32 (43.8)
Never worked/long term unemployed	4 (2.6)	4 (3.0)	4 (5.5)
No father present	26 (17.0)	23 (17.3)	13 (17.8)
Missing data	12 (7.8)	4 (3.0)	-
NSSC(Mother)			
Higher occupations	31 (20.3)	25 (18.8)	13 (17.8)
Intermediate occupations	18 (11.8)	18 (13.5)	12 (16.4)
Lower occupations	66 (43.1)	62 (46.6)	37 (50.7)
Never worked/long term unemployed	26 (17.0)	24 (18.0)	11 (15.1)

Missing data	12 (7.8)	4 (3.0)	-
Mother's education			
School to 16 (no qualifications / certificates below 'A' level)	53 (34.6)	50 (37.6)	23 (31.5)
'A' levels	61 (39.9)	55 (41.4)	35 (47.9)
University Degree/Post-graduate qualification	27 (17.7)	24 (18.0)	15 (20.5)
Missing data	12 (7.8)	4 (3.0)	-

*NSSC: National Statistics Social Class

4. Genotyping

The present study examines whether the effect of artificial food colors (AFCs) on hyperactivity is moderated by genetic difference between children. Such an investigation may help to identify which processes mediate the effect of AFCs on hyperactivity.

Genetic polymorphisms were selected from the dopamine (catechol-o-methyltransferase) and adrenergic (adrenergic receptor alpha 2A) neurotransmitter systems since these have previously been implicated in ADHD. Since there is also a suggestion that histamine may be involved in the effects of AFCs, genetic polymorphisms from this system were also included.

Given the difficulties in genotyping variable number tandem repeat (VNTR) polymorphisms using DNA from cheek cells (the only means of access to DNA for general population samples of children) polymorphisms were selected for the present analysis that were single nucleotide polymorphisms (SNPs). Consequently results are presented here for two SNPs in the histamine N-methyltranferase gene, HNMT Thr105Ile) and HNMT T939C), for one SNP in the dopamine gene, catechol-o-methyltransferase (COMT Val108Met), and one SNP in the adrenergic neurotransmitter system adrenergic receptor, alpha 2A (ADRA2A C1292G).

Table 4.1 presents the distribution of alleles present and absent for each of the genetic polymorphisms outlined both for all children and by gender.

Table 4.1: Genotyping by Gender

	HNMT Thr105Ile %	HNMT T939C %	COMT Val108Met %	ADRA2A C1291G %
Total Sample n = 153	105Ile present = 38 (24.8) 105Ile absent = 97 (63.4) Missing = 18 (11.8)	939C present = 51 (33.3) 939C absent = 81 (52.9) Missing = 21 (13.7)	Val108 present =100 (65.4) Val108 absent = 34 (22.2) Missing = 19 (12.4)	1291G present = 70 (45.8) 1291G absent = 63 (41.2) Missing = 20 (13.1)
Boys n=79	105Ile present = 22 (27.8) 105Ile absent = 46 (58.2) Missing = 11 (13.9)	939C present = 30 (38.0) 939C absent = 36 (45.6) Missing = 13 (16.5)	Val108 present =54 (68.4) Val108 absent = 13 (16.5) Missing = 12 (15.2)	1291G present = 32 (40.5) 1291G absent = 35 (44.3) Missing = 12 (15.2)
Girls n=74	105Ile present = 16 (21.6) 105Ile absent = 51 (68.9) Missing = 7 (9.5)	939C present = 21 (28.4) 939C absent = 45 (60.8) Missing = 8 (10.8)	Val108 present = 46 (62.2) Val108 absent = 21 (28.4) Missing = 7 (9.5)	1291G present = 38 (51.4) 1291G absent = 28 (37.8) Missing = 8 (10.8)

5. Consumption of juice

Parents completed a daily diary of juice consumption and compliance with the diet over the study period. At the end of each week of the study, any bottles containing juice not consumed in the previous week were collected, returned to the study office and the contents measured to help validate, where possible, parental reports of juice consumption by children. Consumption of juice remained at an acceptable level over the period of the study for the majority of children. Of those children (n=16) lost to the study, 12 had a mean of 41% consumption in the first challenge week and data was missing for 4 children. Of the children who completed the study, 128/137 (93%) consumed more than two thirds of all drinks with 103 (80%) of these children consuming ≥85% (at least 6 out of 7 daily drinks per week). Only 1 of the remaining 9 children drank less than 50% of placebo and active drinks over the study period. Table 5.1 presents detail of juice consumption over weeks of the study and by challenge.

Table 5.1: Consumption of juice by week and challenge

	Week 2	Week 4	Week 6	Mix A	Mix B	Placebo
	n (%)			n (%)		
≥ 85 %	113 (74)	112 (73)	105 (69)	108 (71)	109 (71)	113 (74)
≥ 80 %	116 (76)	113 (74)	109 (71)	110 (72)	112 (73)	116 (76)
< 60 %	16 (10)	7 (5)	19 (12)	12 (8)	11 (7)	19 (12)
Missing	4 (3)	16 (10)	16 (10)	12 (8)	14 (9)	10 (7)

Table 5.2 below shows that consumption of juice by challenge was similar for both boys and girls.

Table 5.2: Consumption of juice by challenge and gender

	BOYS (n=79)			GIRLS (n=74)		
	Mix A n (%)	Mix B	Placebo	Mix A n (%)	Mix B	Placebo
≥ 85 %	55 (70)	54 (68)	56 (71)	53 (72)	55 (74)	57 (77)
60 to 84%	12 (15)	11 (14)	6 (8)	9 (12)	8 (11)	5 (7)
< 60 %	5 (6)	7 (9)	10 (13)	7 (9)	4 (5)	9 (12)
Missing	7 (9)	7 (9)	7 (9)	5 (7)	7 (10)	3 (4)

Additional analysis also showed that consumption of juice by challenge was similar for younger and older children.

6. Dietary measures

Pre-study diet measures were collected in order to investigate if diet prior to the study moderated the effect of challenge. Pre study dietary measures are based on 24 hour parental recall of consumption of food items containing additives prior to the start of the study. Table 6.1 presents data relating to the distribution of the number of food items containing additives consumed by all children and by gender groups. This shows that the consumption of additives was similar for boys and girls (all children: median 2.0, range 0 – 5; boys: median 2.0, range 0 - 5; girls: median 2.0, range 0 – 5).

Table 6.1: Dietary intake of foods containing additives by gender

No. of food Items	Total Sample n (%)	Boys n (%)	Girls n (%)
0	34 (22)	19 (24)	15 (20)
1	35 (23)	16 (20)	19 (26)
2	32 (21)	16 (20)	16 (22)
3	25 (16)	13 (17)	12 (16)
4	14 (9)	7 (9)	7 (10)
5	2 (1)	1 (1)	1 (1)
Missing	11 (7)	7 (9)	4 (5)

Dietary infractions

In the weekly diary, completed daily, parents also recorded a ‘whoops’ event each time in any one day the child consumed a portion of food containing artificial colours or sodium benzoate. The occurrence of dietary infractions or ‘mistakes’ by children over the study period was low (33% = 0; 31% = 1 to 2; 18.3% = 3 to 4; 17% > 4). Rates did not differ during active and placebo weeks.

Table 6.2 Dietary infractions ('Whoops') by week and challenge

	Over Weeks			Over Challenge		
	whoops in week 2 n(%)	whoops in week 4 n(%)	whoops in week 6 n(%)	whoops in mix A weeks n(%)	whoops in mix B weeks n(%)	whoops in placebo weeks n(%)
TOTAL						
0	95 (62)	90 (59)	100 (65)	94 (61)	99 (65)	92 (60)
1	34 (22)	33 (22)	15 (10)	29 (19)	25 (16)	28 (18)
>1	9 (6)	10 (7)	17 (11)	10 (7)	11 (7)	15 (10)
Missing	15 (10)	20 (13)	21 (14)	20 (13)	18 (12)	18 (12)
BOYS						
0	49 (62)	48 (61)	48 (61)	49 (62)	47 (60)	49 (62)
1	18 (23)	16 (20)	9 (11)	15 (19)	15 (19)	13 (17)
>1	3 (4)	5 (6)	9 (11)	5 (6)	6 (8)	6 (8)
Missing	9 (11)	10 (13)	13 (17)	10 (13)	11 (14)	11 (14)
GIRLS						
0	46 (62)	42 (57)	52 (70)	45 (61)	52 (70)	43 (58)
1	16 (22)	17 (23)	6 (8)	14 (19)	10 (14)	15 (20)
>1	6 (8)	5 (7)	8 (11)	5 (7)	5 (7)	9 (12)
Missing	6 (8)	10 (14)	8 (11)	10 (14)	7 (10)	7 (10)

7. Behaviour Measures

Behaviour screen measures (Teacher and Parent) ADHD Rating Scale – IV (Teacher version: Pre-school) (DuPaul, Power, Anastopoulos et al, 1997; DuPaul, Power, Anastopoulos & Reid, 1998). Teachers completed behaviour screen questionnaires for all children of the target age within participating schools. This measure provides a total score for 18 items (inattention n=9, hyperactive n=9) directly adapted from the ADHD symptom list as specified in the DSM-IV and scored on a scale of 0 (never or rarely) to 3 (very often) to indicate the frequency of occurrence over the past 6 months. This questionnaire is similar in format to the ADHD rating scale (Teacher version) used for older school children but modified for use with pre-school children. Test-retest reliability coefficients were over .90. Concurrent validity with the Conners Teacher Rating Scale Revised ranged from .55 to .87.

ADHD Rating Scale – IV (Home/Parent version: Pre-School) (DuPaul, Anastopoulos ,Power et al, 1994).

This questionnaire completed by all participating parents prior to baseline (Week 0) is similar in format to the ADHD Rating Scale – IV (Teacher version: Pre-school) with reliability coefficients of .80 and over. Concurrent validity with the Conners Parent Rating Scale Revised ranged from .54 to .96

Behaviour weekly measures (Teacher and Parent)

ADHD Rating Scale – IV (Teacher version: Pre-school) (DuPaul, Power, Anastopoulos et al, 1997)

This questionnaire (see Behaviour Screening above) was completed by the teacher to describe the frequency of the specific behaviour displayed *over the past week* for each week of the study (Week 0 [baseline] to Week 6). Only 10 of the 18 items (inattention n=5, hyperactivity-impulsivity n=5) were completed on a weekly basis and a total score obtained.

Weiss-Werry-Peters (WWP) hyperactivity scale (Routh, 1978).

The WWP has been used in a number of studies to assess hyperactivity (Thompson et al 1996; Hayward et al 1998). Interparent agreement has been found to be good (r=0.82) (Mash & Johnston, 1983). Parents were asked to rate changes within their child's behaviour over the previous week and a total score was obtained for the 7 items in the WWP: (1) switching activities; (2) interrupting or talking too much; (3) wriggling; (4) fiddling with objects or own body; (5) restless; (6) always on the go; (7) concentration.

Classroom observation

Classroom Observation Code (COC) (Abikoff & Gittelman, 1985)

The COC has been fully described in the 8yo Report. The COC assesses the occurrence of 12 mutually exclusive behaviours during structured didactic teaching and during periods of independent work under teacher supervision. A full description of the behaviours coded and scored is given in the COC training manual and these include: interference (to others); interference to teacher; off-task; non-compliance; aggression (physical); verbal aggression to teacher; minor motor movement; gross motor – vigorous; out-of-chair and solicitation of teacher. Children within early years settings in the UK are normally not allocated 'tasks' and are generally free to choose to engage in a range of activities and very little 'structured didactic teaching' sessions take place. For the purposes of the present study, therefore, the 'off-task' code was replaced with a 'switching task' score to reflect the number of occasions on which the child switched activities within the observation period and children were observed while engaged in a range of every day activities. Each child was observed for a total duration of 24 minutes each week (3 observation sessions x 8 minutes duration) and a total weekly score was derived from the total score over each session.

8. Data analysis methods

Generating the behaviour Global Hyperactivity Aggregate (GHA)

The weekly scores from the Teacher, Parent and COC measures for each child are standardised to Time 0 at baseline [T0]) for the same measure –

$$\text{Weekly standardised (z) aggregate score} = \frac{\text{(score X – mean X at T0)}}{\text{SD at T0}}$$

The primary outcome measure, the global hyperactivity aggregate (GHA), is an unweighted aggregate of the weekly Teacher, Parent and COC scores. The GHA is calculated only when at least 2 of these aggregated behaviour scores are present for any week, one of which must be the COC observation score.

Use of Mixed Model method

The effect of challenge (Mix A –v- Placebo; Mix B –v- Placebo) on GHA was tested using Mixed Models Analyses. A number of reports now suggest that traditional methods of analysing repeated measures data, such as end-point analysis and univariate/multivariate repeated measures analysis of variance, have a number of disadvantages (Gueorguieva & Krystal, 2004; Mallinckrodt, Watkin Molenberghs & Carrol, 2004;). Mixed effects models offer a number of advantages via greater flexibility and use of all the available data. Such approaches are able to handle unequally spaced observations over time, correlation between repeated measures on the same subject over time through a random effects approach, differing group sizes and missing data, as long as such data is missing at random. The use of a mixed models analysis also allows a random-effects approach to be combined with patterns of variance in the data over time through specification of a covariance pattern model. For example, selection of a more complex *compound symmetry* covariance structure assumes both equal variances at all time points and equal correlation between any two measurements on the same subject; specification of an *autoregressive* structure of intermediate complexity assumes decreasing correlation between pairs of measures with increasing time difference while a less complex *unstructured* covariance model sets no restrictions at all on the covariance structure.

The present study employed a mixed model analysis to examine the effect of challenge on GHA since missing data was viewed as missing at random and not related to any observed change of behaviour except in the case of one child (see Section 2. Participants). Likelihood ratio tests, reduced SE and covariance parameters indicated that an *compound symmetry* model was the best fit model to apply to the data in the present study.

Power calculations

The study was powered to detect differences between the active and placebo periods and accordingly in each case the effects of Mix A and Mix B were compared to the effect of placebo. It was anticipated that the additional controls on placebo effects would result in an effect size that would be smaller than that achieved in the Isle of Wight study[5]. With a sample of 80 children there was 80% power at alpha = .05 to identify an effect size of 0.32 i.e. the magnitude of the difference in GHA mean score changes (in standard deviation units). The latter being somewhat lower than that achieved in the previous study (0.51). There were uncertainties over the number of children and families who would comply with the demands of a 7 week study and so a target of 120 children was set to reduce the impact of attrition on power and this target was eventually exceeded in both age groups.

9. Behaviour data

Behaviour at screen

Figure 9.1 presents the distribution of teacher behaviour ratings over percentile bands for the whole sample (n=153) and the population of 3 year old children in all participating EYS.

There was no significant difference between the proportions of children in each of the percentile ranges in the FABiC sample and in the rest of the EYS population ($\chi^2(4) = 1.60$, p =.809). The FABiC sample therefore had achieved the approximately equal distribution across these percentile ranges aimed for and were representative of the general population in terms of the Teacher score. The percentile ranges for the raw DuPaul Teacher Scale scores were calculated separately for males and females. This resulted in an approximately equal number of males (n = 79) and females (n = 74) in the FABiC sample. There was no significant difference between the proportions of males in the FABiC sample and in the rest of the EYS population (51.6% -v- 51.1%: χ^2 [1] = 0.00, p =.983).

Figure 9.1 Distribution of teacher ratings for all 3/4yo children in participating EYS

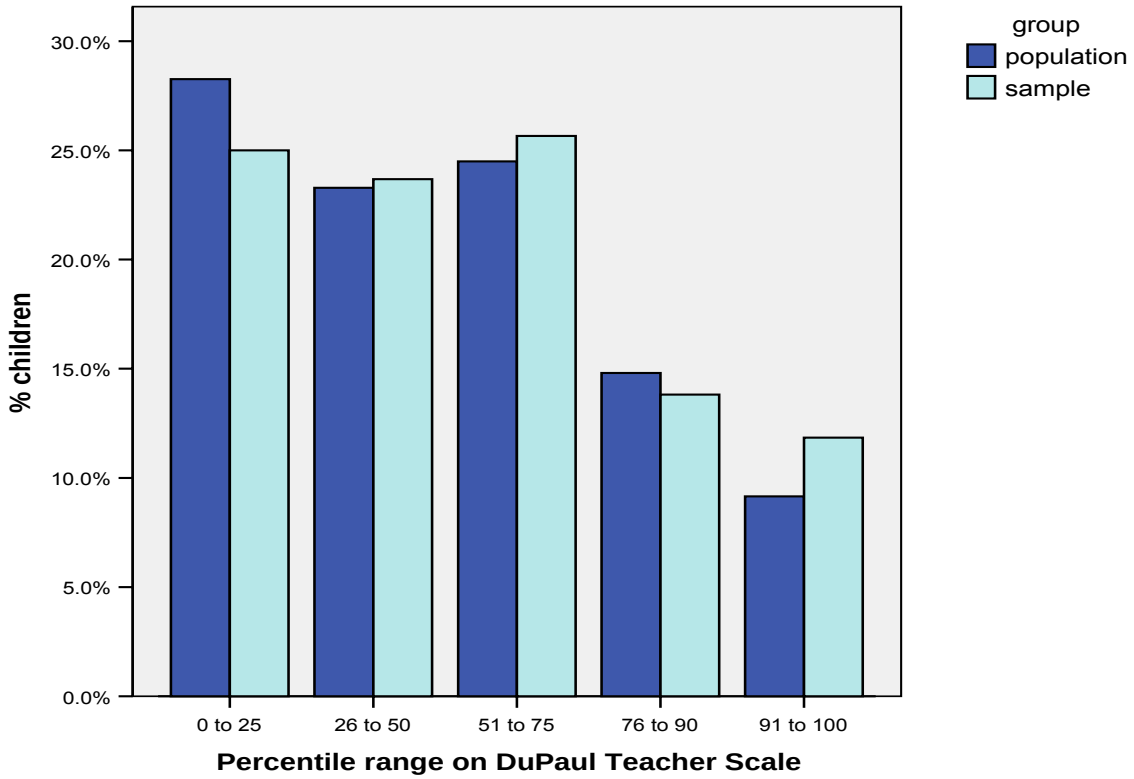


Table 9.1 presents scores (median, inter quartile range [IQR]) for both sample and population by gender. Significant gender differences were found in the total behaviour score for both the population (p<.001) and the sample (p<.001).

Table 9.1: Teacher scores for sample –v- population of 3/4yo children in participating EYS

	Teacher Behaviour Scores					
	Boys		Girls		All	
	Sample (n=79) Median [IQR]	Population (n=380) Median [IQR]	Sample (n=74) Median [IQR]	Population (n=363) Median [IQR]	Sample (n=153) Median [IQR]	Population (n=743) Median [IQR]
Inattention	6 [3 to 11] p= , ns	6 [3 to 12]	3 [0 to 7] p= , ns	3 [0 to 8]	4 [2 to 10] p= , ns	4 [1 to 10]
Hyperactivity	5 [2 to 11] p= , ns	4 [1 to 9]	2 [0 to 7] p= , ns	2 [0 to 5]	3 [1 to 9] p= , ns	3 [1 to 8]
Total behaviour	12 [4 to 22] p= , ns	11 [4 to 20]	6 [2 to 13] p= , ns	5 [1 to 13]	8 [3 to 18] p= , ns	8 [2 to 17]

Behaviour (raw scores) at baseline

Table 9.2 presents behaviour (raw) scores for all children at baseline for each of the behaviour measures employed for all children and by gender

Table 9.2: Behaviour raw scores at baseline x gender

	TOTAL (n = 153)	BOYS (n = 79)	GIRLS (n = 74)
	Median (IQR)	Median (IQR)	Median (IQR)
Teacher	5 (2 to 11)	7 (3 to 13)	4 (1 to 8)
Parent	21 (21 to 22)	21 (21 to 22)	21 (21 to 22)
Classroom Obs	24 (16 to 33)	26 (18 to 37)	21 (13 to 27)

These raw scores were standardised so that an aggregated baseline score (GHA) could be calculated. A comparison of baseline GHA scores showed a significant difference for Gender (t=3.09, p=.002).

Behaviour GHA over challenge period

Tables 9.3 and 9.4 presents the scores over the period of the study standardised to the baseline GHA scores for each measure and for the total aggregated (AGG) GHA. These scores are presented by week and gender (Table 9.3) and by challenge and gender (Table 9.4) .

Table 9.3: Behaviour GHA by Week and Gender

	WEEK 2 Mean (SD)	WEEK 4 Mean (SD)	WEEK 6 Mean (SD)
TOTAL			
Teacher	-.14 (1.01)	-.18 (.88)	-.24 (.78)
Parent	-.20 (1.51)	-.09 (1.78)	-.27 (1.76)
Classroom Obs	.02 (.92)	.01 (.88)	-.19 (.92)
AGG	-.14 (1.07)	-.11 (1.05)	-.32 (1.09)
BOYS (n=79)			
Teacher	.07 (1.09)	-.01 (.93)	-.10 (.76)
Parent	-.25 (1.53)	.12 (1.95)	-.43 (1.67)
Classroom Obs	.23 (.97)	.25 (.90)	-.05 (1.07)
AGG	.03 (1.06)	.14 (1.10)	-.29 (1.13)
GIRLS (n=74)			
Teacher	-.37 (.86)	-.37 (.74)	-.37 (.78)
Parent	-.08 (1.55)	-.40 (1.95)	-.06 (1.75)
Classroom Obs	-.20 (.80)	-.26 (.78)	-.32 (.71)
AGG	-.33 (1.05)	-.38 (.93)	-.35 (1.07)

Table 9.4: Behaviour GHA by Challenge and Gender

	Mix A	Mix B	Placebo
TOTAL			
Teacher	-.18 (.91)	-.14 (.92)	-.23 (.86)
Parent	-.06 (1.66)	-.09 (1.65)	-.41 (1.72)
Class Obs	.02 (.90)	-.08 (.88)	-.08 (.95)
AGG	-.10 (1.04)	-.14 (1.03)	-.33 (1.12)
BOYS (n=79)			
Teacher	.004 (.88)	.05 (1.02)	-.07 (.94)
Parent	-.03 (1.70)	-.18 (1.65)	-.35 (1.85)
Class Obs	.23 (.92)	.07 (1.00)	.17 (1.03)
AGG	.06 (.93)	-.03 (1.16)	-.13 (1.21)
GIRLS (n=74)			
Teacher	-.37 (.90)	-.36 (.73)	-.39 (.73)
Parent	-.09 (1.89)	-.08 (1.78)	-.35 (1.58)
Class Obs	-.20 (.82)	-.25 (.69)	-.33 (.78)
AGG	-.26 (1.14)	-.26 (.87)	-.54 (1.00)

10. Mixed Model Analyses - Likelihood Ratio Tests

Use of MM analysis involves the selection of an appropriate variance model to fit the data. Examination of a range of variance/covariance structures testing the effect of challenge on GHA (≥85% consumption) indicated that the 'unstructured' and 'compound symmetry' matrices were the best fit models. The unstructured matrix is a completely general matrix which imposes no structure on the data while the compound symmetry model assumes a single common variance and a single common correlation between pairs of scores over challenge weeks. While the unstructured matrix had a lower -2LL statistic,(1128.946) compared to the compound symmetry matrix (1133.574), where smallest is best, the use of the 'smaller' compound symmetry matrix involves the calculation of less parameter estimates compared to the unstructured model (2 –v- 6 paramaters). A likelihood ratio test was used to test the null hypothesis that the smaller model provided as good a fit as the unstructured model and this was accepted (p=.400). The compound symmetry model was therefore employed in the MM analyses.

Information Criteria[a]

-2 Restricted Log Likelihood	1128.946
Akaike's Information Criterion (AIC)	1140.946
Hurvich and Tsai's Criterion (AICC)	1141.159
Bozdogan's Criterion (CAIC)	1170.894
Schwarz's Bayesian Criterion (BIC)	1164.894

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: AGG.

Unstructured (6 parameters)

Information Criteria[a]

-2 Restricted Log Likelihood	1133.574
Akaike's Information Criterion (AIC)	1137.574
Hurvich and Tsai's Criterion (AICC)	1137.604
Bozdogan's Criterion (CAIC)	1147.557
Schwarz's Bayesian Criterion (BIC)	1145.557

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: AGG.

Compound Symmetry 2 Parameters)

11. MM main analyses (1)

A total of 153 3 year old children were recruited to enter the study. Of these 16 failed to complete the trial but 149 fulfilled the criteria of having at least one GHA score (Mix A, Mix B or Placebo). The GHA is an aggregate of the week by week teacher and parent ratings, and the Classroom Observation Code (COC) and was calculated when at least two of these behaviour scores were present in any challenge week, one of which had to be the classroom observation score. The analysis in this section just presents the results for the Global Hyperactivity Aggregate (GHA).

Using mixed models to analyse the effects of challenge type, it was found that for these 149 cases there was a significant effect of Mix A (effect size [ES] = .20, p=.039) but not of Mix B (ES = .16) in elevating the GHA compared to placebo (Table 11.1.[1]). When the analysis was restricted to those children (n=133) consuming 85% or more of the juice (Table 11.1.[2]) the effect of Mix A (p=.035) remained significant while the effect of Mix B was non-significant. It should be noted that both effect sizes in this analysis are nevertheless substantial (.24 and .16 for Mix A and Mix B respectively). We also conducted a complete case analysis as a check that the procedure for dealing with missing values in the mixed models method was not producing artefacts in the results. This was limited to the children consuming 85% or more of the drinks and with complete GHA values for Mix A, Mix B and Placebo. These results are presented in Table 11.1.[3]). The pattern of results and the effect sizes for the Mix A vs. Placebo (0.31) and for Mix B vs. Placebo (.19) are similar (in fact slightly larger in the case of Mix A) to those from the mixed model analysis in Table 11.1.[2] i.e. .24 and .16 respectively.

Table 11.1: MM ANALYSIS: effect of Challenge on GHA for 3yo children with ≥0% (1) and ≥85% (2)`juice consumption and complete case analysis (3)

Model	Factor Level	N=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
[1]. (≥0%)										
Challenge (n=149)	Mix A	135	-.10 (1.04)	-.28 to .08	-.11 (.091)	-.29 to .07	A –v- P	.20	.01 to .40	.039
F(2, 257.89)=2.35	Mix B	136	-.14 (1.03)	-.31 to .04	-.16 (.091)	-.34 to .02	B –v- P	.16	-.04 to .35	.113
p=.098	Mix P	132	-.33 (1.12)	-.53 to -.14	-.31 (.092)	-49 to -.13	-	-	-	-
-2LL = 1133.57										
[2]. (≥85%)										
Challenge (n=133)	Mix A	104	-.12 (1.09)	-.33 to .09	-.08 (.103)	-.29 to .12	A –v- P	.24	.02 to .47	.035
F(2, 196.88)=2.32	Mix B	106	-.20 (1.00)	-.40 to -.01	-.17 (.102)	-.37 to .35	B –v- P	.16	-.07 to .38	.173
p=.101	Mix P	106	-.34 (1.10)	-.55 to -.13	-.33 (.102)	-.53 to -.13	-	-	-	
-2LL = 900.63										
[3]. (complete case)										
Challenge (n=73)	Mix A	73	-.14 (1.04)	-.38 to .11	-.14 (.120)	-.38 to .10	Mix A –v- P	.31	.04 to .58	.027
F(2, 144.00)=2.53	Mix B	73	-.26 (1.05)	-.50 to -.01	-.26 (.120)	-.50 to -.02	Mix B –v- P	.19	-.08 to .46	.177
p=.083	Mix P	73	-.44 (.98)	-.67 to -.22	-.44 (.120)	-.68 to -.21	-	-	-	-
-2LL = 613.39										

GHA = Global Hyperactivity Aggregate; Est. Marg.Mean = Estimated Marginal Mean

The challenge periods of central interest were in weeks 2,4 and 6 of the trial. There was no significant tendency for GHA to increase over time (F[4,291.82]=.625, p=.645). To test whether there was any evidence of carry-over effects the scores of the previous active challenge period and baseline were added as factors in the mixed model analysis. No significant effect of the previous active period was found (F[3,219.41]=.830, p=.479). These factors were not included in further analyses.

12. Other possible influencing factors in MM analysis

It was possible that the effects of the additive mixes would be most marked for those already showing higher GHA scores at baseline. The effect of a number of factors on behaviour measured at baseline was investigated (Table 12.1). This indicated that gender and pre-study intake of foods containing additives were related to behaviour at baseline.

Table 12.1: Effect of possible influencing factors on Baseline GHA

Variable	N	All Mean (sd)	Groups Mean (sd)	 Mean (sd)	Test
			Age ≤43m (n=66)	Age >43m (n=67)	
Age Gp	133	-.05 (.68)	-.14 (.69)	.07 (.66)	F(1,131)=3.18, p=.077
			Male (n=67)	Female (n=66)	
Gender	133	-.05 (.68)	.14 (.78)	-.22 (.51)	F(1,131)=9.73, p=.002
			O levels or less (n=50)	A levels or higher (n=79)	
Maternal education level	129	-.06 (.68)	-.02 (.63)	-.08 (.71)	F(1,127)=.215, p=.644
			Pre-study diet		
Pre-study diet	131	-.05 (.68)	-	-	F(1,129)=4.00, p=.048

Preliminary analyses showed that the baseline total GHA behaviour score was related to the challenge GHA and that gender and pre-study diet were individually related to GHA but in no case was there a significant interaction between these factors and challenge type. Boys had significantly higher GHA scores than girls and children with a higher number of foods containing additives in their diet, as measured by parental 24 hour recall prior to the start of the study, had higher GHA scores.

The effect of inclusion of baseline GHA, gender and pre-study in the MM analysis were examined separately. Significance persisted for baseline GHA ($F[1, 130.31]=32.03$, $p=.000$), gender ($F[1, 122.11]=6.94$, $p=.01$) and pre-study diet ($F[1, 119.64]=4.05$, $p=.047$) but in no case was there a significant interaction effect with challenge. However, when both GHA and gender were included in analyses, gender was no longer significant, probably due to the shared variance between these two measures. All subsequent analyses, therefore, included baseline GHA and pre-study diet only.

13. MM main analyses (2)

The MM Analyses examining the effect of challenge on GHA shown in Table 11.1 [1], [2] and [3] was repeated below with baseline GHA and pre-study diet included in the analysis of all children [1], children with at least 85% juice consumption [2] and a complete case analysis [3].

Table 13.1 shows that the significant effect of Mix A on GHA persists and remains significant in all analyses when baseline GHA and pre-study diet are included in analyses. The effect of Mix B on GHA was not significant.

Table 13.1. [1, 2 and 3]: MM Analysis of effect of Challenge on GHA for 3yo children

Mixed Model *	Factor level	N=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
[1]. (≥0%) Challenge (n=140) F(2, 258.52)=2.15 p=.119 -2LL = 1075.45	Mix A	131	-.11 (1.03)	-.29 to .07	-.12 (.086)	-.29 to .05	Mix A –v- P	.19	.001 to .39	.048
	Mix B	134	-.14 (1.03)	-.32 to .03	-.16 (.086)	-.33 to .007	Mix B –v- P	.15	-.04 to .34	.128
	Mix P	129	-.32 (1.11)	-.51 to -.12	-.31 (.087)	-.48 to -.14	-	-	-	-
[2]. (≥85%) Challenge (n=131) F(2, 202.09)=2.49 p=.085 -2LL = 871.79	Mix A	102	-.12 (1.09)	-.33 to .10	-.09 (.098)	-.28 to .10	Mix A –v- P	.25	.03 to .48	.028
	Mix B	106	-.20 (1.00)	-.40 to -.01	-.19 (.096)	-.37 to .004	Mix B –v- P	.16	-.07 to .38	.168
	Mix P	106	-.34 (1.10)	-.55 to -.13	-.34 (.096)	-.53 to -.15	-	-	-	-
[3]. Challenge (n=73) F(2, 144.00)=2.53 p=.083 -2LL = 609.43	Mix A	73	-.14 (1.04)	-.38 to .11	-.14 (.117)	-.37 to .09	Mix A –v- P	.31	.04 to .58	.027
	Mix B	73	-.26 (1.05)	-.50 to -.01	-.26 (.117)	-.49 to -.03	Mix B –v- P	.19	-.08 to .46	.177
	Mix P	73	-.44 (.98)	-.67 to -.22	-.44 (.117)	-.68 to -.21	-	-	-	-

Effect of Challenge on GHA with Baseline GHA and Pre-study diet included in Model

14. MM analyses using all covariates (results report in Lancet (2007) paper)

For the final MM analyses the effect of challenge on GHA was examined again but now with all potential confounds controlled. The results are shown in Table 14.1 for all children, children with at least 85% juice consumption and a complete case analysis. The format of this table is slightly simplified to accommodate parameters for all the covariates.

The results show that for all three analyses there was a statistically significant greater GHA score when challenged with Mix A than with Placebo. For Mix B the effect size of 0.16 or greater just failed to reach significance at the $p<.05$ level.

Table 14.1. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	Whole sample n=140	≥85% consumption n=130	Complete case n=73
Model 1			
Intercept	-.31 [-.49 to -.13] **	-.33 [-.53 to -.13] ***	-.44 [-.68 to -.21] ***
Challenge type			
Mix A –v- placebo	.20 [0.01 to .40] *	.24 [.02 to .47] *	.31 [.04 to .58] *
Mix B –v- placebo	.16 [-.04 to .35]	.16 [-.07 to .38]	.19 [-.08 to .46]
Model 2			
Intercept	-.54 [-.89 to -.18] **	-.51 [-.92 to -.11]	-.58 [-1.08 to -.09]*
Challenge type			
Mix A –v- placebo	.20 [.01 to .39] *	.28 [.05 to .51] *	.32 [.05 to .60] *
Mix B –v- placebo	.17 [-.03 to .36]	.19 [-.04 to .41]	.21 [-.06 to .48]
Week of study			
Wk 2 –v- Wk 6	.15 [-.05 to .34]	.15 [-.08 to .38]	.19 [-.08 to .46]
Wk 4 –v- Wk 6	.17 [-.03 to .36]	.23 [.00 to .46] *	.19 [-.09 to .46]
Gender	.18 [-.10 to .45]	.22 [-.07 to .51]	.05 [-.31 to .40]
Baseline GHA score	.46 [.26 to .66] ***	.54 [.31 to .76] ***	.36 [.06 to .66] *
Pre-trial diet	.08 [-.02 to .19]	.07 [-.04 to .18]	.09 [-.04 to .23]
Maternal education level	-.01 [-.29 to .28]	-.04 [-.34 to .26]	-.03 [-.41 to .35]
Maternal social class	.15 [-.44 to .13]	-.23 [-.53 to .08]	-.21 [-.58 to .16]

* p<.05, ** p<.01, *** p<.001

15. Genotype MM analyses (1)

The following analyses investigate whether the effect of artificial food colors (AFCs) on hyperactivity is moderated by genetic difference between children. Consequently results are presented here for the following genetic polymorphisms: HNMT Thr105Ile, HNMT T939C, COMT Val108Met and ADRA2A c1291G.

Similar procedures were employed to examine the effect of each of the genetic polymorphisms outlined in Section 4. Their relationship to Baseline GHA was first examined prior to examining their individual effect and interaction with Challenge on GHA in the MM analyses. Table 15.1 indicates no significant effect of these polymorphisms on baseline GHA although the effect of HNMT Thr105Ile and HNMT T939C do approach significance.

Table 15.1: Baseline analyses of Genotype

Moderator	n	All Mean (sd)	Groups Mean (sd)	 Mean (sd)	Test
			105Ile present (n=36)	105Ile absent (n=92)	
HNMT Thr 105Ile	128	-.05 (.68)	-.11 (.65)	-.12 (.68)	F(1,126)=2.88, p=.092
			939C present (n=47)	939C absent (n=78)	
HNMT T939C	125	-.04 (.68)	.10 (.60)	-.13 (.71)	F(1,123)=3.41, p=.067
			Val108 present (n=95)	Val108 absent (n=32)	
COMT Val108Met	127	-.05 (.67)	-.06 (.68)	.002 (.65)	F(1,125)=.21, p=.646
			1291G present (n=66)	1291G absent (n=60)	
ADRA2A C1291G	126	-.06 (.68)	-.13 (.68)	.02 (.68)	F(1,124)=1.56, p=.213

Their effect on GHA in the MM analyses is presented in Tables 15.2 to 15.5 below. The analyses in Table 15.2 (HNMT Thr105Ile) and 15.3 (HNMT T939C) show significant main effects (p=..005 and p=.012 respectively) and Challenge x Genotype interactions which approach significance for HNMT Thr105Ile (Mix A: p=.05; Mix B: p=.091) but less so for HNMT T939C (Mix A: p=.097; Mix B: p=.290).

Table 15.2: MM Analysis of effect of Challenge and Genotype (HNMT Thr105Ile) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=128) -2LL = 853.23	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 194.91)=.713 p=.339	Mix A	100	-.12 (1.10)	-.34 to .10	-.06 (.114)	-.29 to .16	Mix A –v- P	.36	.10 to .62	.008
	Mix B	104	-.17 (.97)	-.36 to .02	-.09 (.119)	-.33 to .14	Mix B –v- P	.30	.04 to .55	.023
	Mix P	103	-.33 (1.09)	-.54 to -.12	-.17 (.115)	-.40 to .06	-	-	-	-
(≥85%) HNMT Thr105Ile F(1, 121.42)=3.58 p=.061	105Ile present	36	-.02 (1.27)	-.30 to .27	.06 (.150)	-.24 to .36	105ile Pre v Ab	.65	.20 to 1.11	.005
	105Ile absent	92	-.27 (.96)	-.40 to -.15	-.28 (.092)	-.46 to -.09				
(≥85%) Challenge x HNMT Thr105Ile - 105Ile present F(2, 194.91)=2.30 p=.103	Mix A	29	-.06 (1.217)	-.50 to .39	.01 (.192)	-.37 to .39	)	105Ile present –v- 105Ile absent: Mix A –v- P: -.51 (-1.02 to -.001), p=.05 Mix B –v- P: -.45 (-.97 to .07), p=.091		
	Mix B	24	-.03 (1.26)	-.56 to .50	.01 (.207)	-.40 to .42	)			
	Mix P	27	.04 (1.43)	-.53 to .60	.16 (.198)	-.23 to .55	)			
							)			
(≥85%) Challenge x HNMT Thr105Ile - 105Ile absent	Mix A	71	-.14 (1.08)	-.34 to .11	-.14 (.122)	-.38 to .11	)			
	Mix B	80	-.21 (.88)	-.41 to -.02	-.20 (.117)	-.43 to .03	)			
	Mix P	76	-.46 (.91)	-.67 to -.25	-.49 (.119)	-.73 to -.26	)			

Table 15.3: MM Analysis of effect of Challenge and Genotype (HNMT T939C) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=125) -2LL = 851.21	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge										
	Mix A	97	-.12 (1.12)	-.35 to .10	-.08 (.108)	-.29 to .13	Mix A –v- P	.38	.08 to .67	.013
F(2, 186.06)=1.23	Mix B	102	-.16 (.98)	-.35 to .03	-.09 (.109)	-.31 to .12	Mix B –v- P	.29	.002 to .57	.049
p=.295	Mix P	100	-.33 (1.10)	-.55 to -.11	-.25 (.108)	-.47 to -.04	-	-	-	-
(≥85%) HNMT T939C	939C present	47	-.03 (1.15)	-.25 to .19	.02 (.131)	-.24 to .28	939C Pre v Ab	.55	.12 to .97	.012
F(1, 114.77)=3.83	939C absent	78	-.31 (1.00)	-.45 to -.16	-.30 (.101)	-.50 to -.10				
p=.053										
(≥85%) Challenge x HNMT T939C – 939C present							)			
	Mix A	38	-.03 (1.07)	-.38 to .32	-.01 (.169)	-.34 to .32	)			
	Mix B	34	.003 (1.13)	-.39 to .40	.05 (.177)	-.30 to .40	)			
F(2, 184.06)=1.43	Mix P	37	-.05 (1.28)	-.48 to .38	.02 (.171)	-.32 to .36	)	939C present –v- 939C absent:		
P=.242							)	Mix A –v- P: -.41 (-.89 to .07),		
(≥85%) Challenge x HNMT T939C - 939C absent							)	p=.097		
	Mix A	59	-.18 (1.15)	-.48 to .12	-.15 (.135)	-.41 to .12	)	Mix B –v- P: -.26 (-.74 to .22),		
								p=.290		
	Mix B	68	-.24 (.89)	-.46 to -.03	-.24 (.128)	-.49 to .02	)			
	Mix P	63	-.31 (1.00)	-.73 to -.26	-.52 (.132)	-.78 to -.27	)			

Table 15.4: MM Analysis of effect of Challenge and Genotype (COMT Val108Met) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=127) -2LL = 856.04	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 190.32)=2.13 p=.122	Mix A	99	-.14 (1.08)	-.36 to .07	-.01 (.118)	-.24 to .23	Mix A –v- P	.42	-.03 to .88	.067
	Mix B	103	-.17 (.98)	-.36 to .02	-.12 (.120)	-.36 to .11	Mix B –v- P	.11	-.36 to .58	.645
	Mix P	102	-.33 (1.09)	-.55 to -.12	-.28 (.120)	-.52 to -.04	-	-	-	-
(≥85%) COMT Val108Met F(1, 117.94)=1.65 p=.201	Val108 present	95	-.27 (1.06)	-.41 to -.13	-.25 (.090)	-.43 to -.07	Val108 Pre v Ab	-.16	-.64 to .31	.493
	Val108 absent	32	-.09 (.87)	-.29 to .11	-.02 (.158)	-.33 to .29				
(≥85%) Challenge x COMT Val108Met - Val108 present F(2, 190.32)=1.19 p=.305	Mix A	73	-.27 (1.10)	-.53 to -.02	-.24 (.120)	-.47 to -.002	))			
	Mix B	79	-.16 (1.02)	-.39 to .07	-.16 (.116)	-.39 to .07	)			
	Mix P	78	-.34 (1.18)	-.61 to -.07	-.36 (.117)	-.59 to -.13	))	Val108 present –v- Val108 absent:		
(≥85%) Challenge x COMT Val108Met - Val108 absent	Mix A	26	.21 (.94)	-.17 to .59	.23 (.203)	-.17 to .63	)	Mix A –v- P: -.30 (-.82 to .22), p=.261		
	Mix B	24	-.20 (.86)	-.56 to .16	-.09 (.209)	-.50 to .32	)	Mix B –v- P: .09 (-.44 to .63), p=.730		
	Mix P	24	-.30 (.73)	-.61 to .01	-.20 (.210)	-.61 to .22	)			

Table 15.5: MM Analysis of effect of Challenge and Genotype (ADRA2A C1291G) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=126) -2LL = 846.41	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 186.06)=2.60 p=.077	Mix A	98	-.10 (1.09)	-.32 to .12	-.07 (.104)	-.27 to .14	Mix A –v- P	.25	-.08 to .57	.132
	Mix B	103	-.16 (.98)	-.36 to .03	-.13 (.103)	-.34 to .07	Mix B –v- P	.09	-.24 to .42	.588
	Mix P	102	-.33 (1.09)	-.55 to -.12	-.32 (.103)	-.52 to -.12	-	-	-	-
(≥85%) ADRA2A C1291G F(1, 113.81)=2.71 p=.103	1291G present	66	-.31 (1.23)	-.50 to -.12	-.30 (.109)	-.52 to -.09	1292G Pre v Ab	-.33	-.73 to .08	.114
	1291G absent	60	-.08 (.81)	-.21 to .05	-.04 (.114)	-.27 to .18				
(≥85%) Challenge x ADRA2A C1291G -1291G present F(2, 186.06)=.421 p=.657	Mix A	49	-.26 (1.22)	-.61 to .09	-.22 (.147)	-.51 to .07	))			
	Mix B	57	-.22 (1.12)	-.52 to .08	-.20 (.139)	-.48 to .07	)			
	Mix P	53	-.45 (1.35)	-.82 to -.07	-.48 (.142)	-.76 to -.20	)	1291G present –v- 1291G absent:		
							)	Mix A –v- P: .01 (-.45 to .47), p=.962		
(≥85%) Challenge x ADRA2A C1291G - 1291G absent	Mix A	49	.07 (.94)	-.20 to .34	.09 (.148)	-.20 to .39	)			
	Mix B	46	-.07 (.77)	-.32 to .14	-.07 (.152)	-.37 to .24	)	Mix B –v- P: .19 (-.26 to .64), p=.413		
	Mix P	49	-.21 (.69)	-.41 to -.01	-.16 (.149)	-.45 to .14	)			

16. Genotype MM analyses (2)

With the inclusion of Baseline GHA in the MM main analyses which did not include genotype, the addition of Gender to the model did not have any added effect since Baseline GHA and the effect of Gender on behaviour would have shared variance. However, gender has been shown to moderate the effects of genotype in a number of ways over a number of conditions in children. For this reason the above Genotype analyses were re-examined with the inclusion of Gender and a Genotype by Gender interaction term. This was carried out in order to check if the interaction term accounted for variance over and above that accounted for by Baseline GHA in these Genotype analyses. Analyses were only carried out for HNMT Thr105Ile and HNMT T939C where a significant relationship between the polymorphism and GHA was evident. No further analyses have been carried out for ADRA2A C1291G or COMT Val108Met.

With further inclusion of Gender and the Genotype by Gender interaction, the analysis in Table 16.1 shows that the effect of both Mix A and Mix B on GHA are significant, the difference between HNMT 105Ile present and absent alleles is significantly greater for boys compared to girls (p=.021) and the difference between 105Ile present –v- absent approaches significance for Mix A compared to Mix P (p=.053) with 105Ile absent (Mix A –v- P) showing significantly higher scores. No significant effect was found for the Mix B –v- P (p=.090) comparison.

Similarly, Analysis 16.2 shows a significant effect of Mix A on GHA (p=.013) and for Mix B (p=.049), the difference between HNMT 939C present and absent alleles is significantly greater for boys compared to girls (p=.025) and the difference between 939C present –v- absent approaches significance for Mix A compared to Mix P (p=.097).

Table 16.1: MM Analysis of effect of Challenge, Genotype (HNMT Thr105Ile), Gender, Challenge x HNMT Thr105Ile, Gender x HNMT Thr105Ile on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=128) -2LL = 854.33	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 198.05)=.347 p=.707	Mix A	100	-.12 (1.10)	-.34 to .10	-.09 (.111)	-.31 to .13	Mix A –v- P	.36	.09 to .62	.008
	Mix B	104	-.17 (.97)	-.36 to .02	-.13 (.116)	-.35 to .11	Mix B –v- P	.30	.04 to .55	.023
	Mix P	103	-.33 (1.09)	-.54 to -.12	-.20 (.113)	-.42 to .03	-	-	-	-
(≥85%) Gender F(1, 118.81)=12.96 p=.000	Male	64	-.01 (1.07)	-.18 to .16	.17 (.116)	-.06 to .40	Male v Female	.21	-.14 to .56	.232
	Female	64	-.40 (1.01)	-.56 to -.24	-.44 (.124)	-.69 to -.20				
(≥85%) HNMT Thr105Ile F(1, 119.21)=2.45 p=.120	105Ile present	36	-.02 (1.27)	-.30 to .27	-.004 (.145)	-.29 to .28	105Ile Pre v Ab	.18	-.39 to .75	.524
	105Ile absent	92	-.27 (.96)	-.40 to -.15	-.27 (.088)	-.45 to -.10				
(≥85%) HNMT Thr105Ile x Gender (Male) F(1, 118.81)=5.52 p=.021	105Ile present	21	.43 (1.13)	.08 to .79	.50 (.193)	.12 to .88	))))))	105Ile Present -v- absent: Male –v- Female: .80, .12 to 1.47, **p= .021**		
	105Ile absent	43	-.18 (.99)	-.37 to .01	-.16 (.128)	-.42 to .09				
(≥85%) HNMT	105Ile	15	-.51 (1.25)	-.92 to -.10	-.51 (.216)	-.94 to -.08	)			

Thr105Ile x Gender (Female)	present 105Ile absent	49	-.36 (.93)	-.53 to -.19	-.38 (.122)	-.62 to -.14		
(≥85%) Challenge x HNMT Thr105Ile -105Ile present F(2, 198.05)=2.26 p=.107	Mix A	29	-.06 (1.217)	-.50 to .39	-.05 (.188)	-.42 to .32	)	105Ile present –v- 105Ile absent: Mix A –v- P -.57, -1.10 to -.03, **p=.037** Mix B –v- P -.37, -.88 to .15, p=.160
	Mix B	24	-.03 (1.26)	-.56 to .50	-.06 (.203)	-.46 to .34	)	
	Mix P	27	.04 (1.43)	-.53 to .60	.10 (.193)	-.29 to .48	)	
							)	
							)	
(≥85%) Challenge x HNMT Thr105Ile - 105Ile absent	Mix A	71	-.14 (1.08)	-.34 to .11	-.13 (.188)	-.37 to .10		
	Mix B	80	-.21 (.88)	-.41 to -.02	-.19 (.114)	-.42 to .03	)	
	Mix P	76	-.46 (.91)	-.67 to -.25	-.49 (.116)	-.72 to -.26	)	

Table 16.2: MM Analysis of effect of Challenge and Genotype (HNMT T939C) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=125) -2LL = 842.17	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 188.45)=1.20 p=.305	Mix A	97	-.12 (1.12)	-.35 to .10	-.12 (.107)	-.33 to .09	Mix A –v- P	.37	.08 to .69	.013
	Mix B	102	-.16 (.98)	-.35 to .03	-.14 (.108)	-.35 to .08	Mix B –v- P	.29	.001 to .57	.049
	Mix P	100	-.33 (1.10)	-.55 to -.11	-.29 (.107)	-.50 to -.08	-	-	-	-
(≥85%) Gender F(1, 111.46)=9.39 p=.003	Male	62	-.01 (1.08)	-.19 to .16	.07 (.110)	-.15 to .28	Male v Female	.13	-.26 to .52	.512
	Female	63	-.39 (1.02)	-.55 to -.23	-.43 (.118)	-.66 to -.19				
(≥85%) HNMT T939C F(1, 111.67)=1.98 p=.162	939C present	47	-.03 (1.15)	-.25 to .19	-.07 (.128)	-.32 to .19	939C Pre v Ab	.08	-.46 to .63	.757
	939C absent	78	-.31 (1.00)	-.45 to -.16	-.30 (.098)	-.49 to -.10				
(≥85%) HNMT T939C x Gender (Male) F(1, 111.46)=5.14 p=.025	939C present	29	.31 (1.03)	.05 to .57	.36 (.163)	.04 to .68	))))))	939C Present -v- 939C absent: Male –v- Female: .73, .09 to 1.37, **p= .025**		
	939C absent	33	-.26 (1.06)	-.49 to -.03	-.23 (.148)	-.52 to .06				
(≥85%)	C present	18	-.50 (1.15)	-.84 to -.15	-.50 (.198)	-.89 to -.10				

HNMT T939C x Gender (Female)	C absent	45	-.34 (.96)	-.53 to -.16	-.36 (.128)	-.61 to -.10		
(≥85%) Challenge x HNMT T939C - 939C present F(2, 188.45)=1.43 p=.241	Mix A	38	-.03 (1.07)	-.38 to .32	-.10 (.168)	-.43 to .23	)	939C present –v- 939C absent: Mix A –v- P -.41, -.89 to .07, p=.097
	Mix B	34	.003 (1.13)	-.39 to .40	-.04 (.176)	-.39 to .30	)	
	Mix P	37	-.05 (1.28)	-.48 to .38	-.06 (.169)	-.40 to .27	)	
(≥85%) Challenge x HNMT T939C - 939C absent	Mix A	59	-.18 (1.15)	-.48 to .12	-.14 (.133)	-.40 to .12	)	Mix B –v- P -.26, -.74 to .22, p=.281
	Mix B	68	-.24 (.89)	-.46 to -.03	-.23 (.126)	-.48 to .02	)	
	Mix P	63	-.31 (1.00)	-.73 to -.26	-.51 (.130)	-.77 to -.26	)	

17. Genotype MM analyses (3)

These prior Genotype analyses indicate that the HNMT Thr105Ile and possibly the HNMT T939C polymorphisms play a possible moderating role in the relationship between Challenge and GHA. Certainly in the case of the HNMT Thr105Ile polymorphism, those children with the 105Ile absent genotype may be more vulnerable to the effects of additives in their food. Further analyses are presented below for each of these polymorphisms, but including both Baseline GHA and Pre-study diet and based only on children with ≥85% juice consumption. Table 17.1 shows that the inclusion of Baseline GHA and Pre-study Diet has made no difference to findings. The effect of both Mix A and Mix B on GHA remain significant (p=.006 and p.030 respectively), the difference between HNMT 105Ile present and absent genotypes is significantly greater for boys compared to girls (p=.019) and the difference between 105Ile present –v- absent is significantly greater for Mix A compared to Mix P (p=.042) with 105Ile absent (Mix A –v- P) showing significantly higher scores. No significant difference was found between the Thr105Ile genotypes for the Mix B –v- P (p=.116) comparison. Baseline GHA remains significant in this analysis (ES: .51, .29 to .73, p=.000) while the effect of Pre-study Diet is not significant (ES: .05, -.06 to .15, p=.362). These values are not shown in Analysis 17.1.

Similarly, Table 17.2 shows the effect of Mix A on GHA is again significant (p=.010) but only approaches significance for Mix B (p=.063), the difference between HNMT T939C C present and absent genotypes is significantly greater for boys compared to girls (p=.029) and the difference between 939C present –v- absent only approaches significance for Mix A compared to Mix P (p=.076) with 939C absent (Mix A –v- P) showing increased scores. No significant effect was found for the Mix B –v- P (p=.352) comparison. Baseline GHA again remains significant (ES: .52, .30 to .74, p=.000) while Pre-study Diet remains non-significant (ES: .05, -.06 to .15, p=.399).

These Genotype analyses indicate that the HNMT Thr105Ile polymorphism, and possibly HNMT T939c polymorphism, may play a possible moderating role in the relationship between Challenge and GHA. Certainly in the case of the HNMT Thr105Ile polymorphism, those children with the 105Ile absent genotype may be more vulnerable to the effects of additives in their food compared to those with the allele present.

Table 17.1: MM Analysis of effect of Challenge, Genotype (HNMT Thr105Ile), Gender, Challenge x HNMT Thr105Ile, Gender x HNMT Thr105Ile, Baseline GHA (not shown) and Pre-study Diet (not shown) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=128) -2LL = 837.71	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 202.13)=.358 p=.700	Mix A	100	-.12 (1.10)	-.34 to .10	-.12 (.106)	-.33 to .09	Mix A –v- P	.37	.11 to .63	.006
	Mix B	104	-.17 (.97)	-.36 to .02	-.15 (.111)	-.37 to .07	Mix B –v- P	.28	.03 to .54	.030
	Mix P	103	-.33 (1.09)	-.54 to -.12	-.23 (.107)	-.44 to -.02	-	-	-	-
(≥85%) Gender F(1, 115.37)=7.13 p=.009	Male	64	-.01 (1.07)	-.18 to .16	.05 (.109)	-.17 to .26	Male v Female	.06	-.27 to .38	.728
	Female	64	-.40 (1.01)	-.56 to -.24	-.38 (.114)	-.61 to -.15				
(≥85%) HNMT Thr105Ile F(1, 117.87)=1.10 p=.296	105Ile present	36	-.02 (1.27)	-.30 to .27	-.08 (.134)	-.35 to .18	105Ile Pre v Ab	.11	-.43 to .64	.692
	105Ile absent	92	-.27 (.96)	-.40 to -.15	-.25 (.081)	-.41 to -.09				
(≥85%) HNMT Thr105Ile x Gender (Male) F(1, 117.52)=5.63 p=.019	105Ile present	21	.43 (1.13)	.08 to .79	.31 (.181)	-.05 to .67	))))))	105Ile Present -v- 105Ile absent: Male –v- Female: .74, .12 to 1.36, **p= .019**		
	105Ile absent	43	-.18 (.99)	-.37 to .01	-.22 (.118)	-.45 to .01				
(≥85%) HNMT Thr105Ile x	105Ile present	15	-.51 (1.25)	-.92 to -.10	-.48 (.198)	-.88 to -.09				

Gender (Female)	105Ile absent	49	-.36 (.93)	-.53 to -.19	-.28 (.113)	-.50 to -.05		
(≥85%) Challenge x HNMT Thr105Ile – 105Ile present F(2, 202.08)=2.31 p=.102	Mix A	29	-.06 (1.217)	-.50 to .39	-.15 (.179)	-.50 to .21	)	105Ile present –v- absent: Mix A –v- P -.52, -1.03 to -.02, **p=.042** Mix B –v- P -.41, -.93 to .10, p=.116
	Mix B	24	-.03 (1.26)	-.56 to .50	-.12 (.194)	-.50 to .26	)	
	Mix P	27	.04 (1.43)	-.53 to .60	.01 (.185)	-.35 to .38	)	
							)	
							)	
(≥85%) Challenge x HNMT Thr105Ile - 105Ile absent	Mix A	71	-.14 (1.08)	-.34 to .11	-.10 (.114)	-.32 to .13		
	Mix B	80	-.21 (.88)	-.41 to -.02	-.18 (.108)	-.40 to .03	)	
	Mix P	76	-.46 (.91)	-.67 to -.25	-.47 (.110)	-.68 to -.25	)	

Table 17.2: MM Analysis of effect of Challenge and Genotype (HNMT T939C), Baseline GHA (not shown) and Pre-study Diet (not shown) on GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=125) -2LL = 825.90	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 192.71)=1.24 p=.293	Mix A	97	-.12 (1.12)	-.35 to .10	-.14 (.102)	-.34 to .06	Mix A –v- P	.39	.09 to .68	.010
	Mix B	102	-.16 (.98)	-.35 to .03	-.15 (.103)	-.36 to .05	Mix B –v- P	.27	-.01 to .55	.063
	Mix P	100	-.33 (1.10)	-.55 to -.11	-.31 (.102)	-.51 to -.11	-	-	-	-
(≥85%) Gender F(1, 108.63)=4.34 p=.039	Male	62	-.01 (1.08)	-.19 to .16	-.04 (.103)	-.25 to .16	Male v Female	-.01	-.37 to .35	.954
	Female	63	-.39 (1.02)	-.55 to -.23	-.36 (.110)	-.58 to -.14				
(≥85%) HNMT T939C F(1, 111.12)=.673 p=.414	939C present	47	-.03 (1.15)	-.25 to .19	-.14 (.119)	-.37 to .10	939C Pre v Ab	.01	-.49 to .52	.957
	939C absent	78	-.31 (1.00)	-.45 to -.16	-.26 (.090)	-.44 to -.08				
(≥85%) HNMT T939C x Gender (Male) F(1, 110.03)=4.88 p=.029	939C present	29	.31 (1.03)	.05 to .57	.19 (.154)	-.12 to .49	))))))))	939C Present -v- 939C absent: Male –v- Female: .66, .07 to 1.25, **p= .029**		
	939C absent	33	-.26 (1.06)	-.49 to -.03	-.27 (.136)	-.54 to .003				
(≥85%)	939C	18	-.50 (1.15)	-.84 to -.15	-.46 (.182)	-.82 to -.10				

HNMT T939C x Gender (Female)	present 939C absent	45	-.34 (.96)	-.53 to -.16	-.26 (.120)	-.49 to -.02		
(≥85%) Challenge x HNMT T939C - 939C present F(2, 192.54)=1.59 p=.206	Mix A	38	-.03 (1.07)	-.38 to .32	-.18 (.160)	-.50 to .13	)	939C present –v- absent: Mix A –v- P -.43, -.91 to .05, **p=.076** Mix B –v- P -.23, -.70 to .25, p=.352
	Mix B	34	.003 (1.13)	-.39 to .40	-.10 (.168)	-.43 to .23	)	
	Mix P	37	-.05 (1.28)	-.48 to .38	-.14 (.162)	-.46 to .18	)	
(≥85%) Challenge x HNMT T939C - 939C absent	Mix A	59	-.18 (1.15)	-.48 to .12	-.09 (.127)	-.34 to .16	)	
	Mix B	68	-.24 (.89)	-.46 to -.03	-.21 (.120)	-.45 to .02	)	
	Mix P	63	-.31 (1.00)	-.73 to -.26	-.48 (.123)	-.72 to -.24	)	

18. Genotype MM analyses using all covariates

The analyses of the moderating effects of genotype were run again but on this occasion using all the potential confound that had been included in the analyses in Section 14. The results are presented in the same reduced form as was adopted in Section 14 to provided parameters on all the covariates. The results are shown in Table 18.1 to Table 18.2 for children with at least 85% juice consumption for the HNMT gene polymorphisms.

It can be seen in Table 18.1 that with the full set of covariates that the GHA is higher for HNMT 105Ile present genotype (ES: 51, .08 to .95, p=.021). The adverse effect of Mix A compared to Placebo is greater for the HNMT 105Ile absent genotype than for the 105Ile present genotype (ES: -.53, -1.04 to -.02, p=.041). The effects in the same direction do not reach significance for Mix B compared to Placebo (ES: -.40, -.92 to .12, p=.134). In Table 18.2 the pattern of results for HNMT T939C is that for 939C absent the effects of Mix A compared to Placebo are greater than for 939C present but this just fails to reach significance (ES: -.46, -.94 to .02,p=.061). There is no evidence that HNMT T939C moderates the effects of Mix B (ES: -.23, -.72 to .25, p=.338).

Table 18.1. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (HNMT Thr105Ile) and Challenge type by Genotype interaction for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	≥85% consumption
Model 1 Intercept	-.49 (-.73 to -.26), p<.001
Challenge type	
Mix A –v- placebo	.36 (.10 to .62), p=.008
Mix B –v- placebo	.30 (.04 to .55), p=.023
Genotype – HNMT Thr105Ile 105Ile present –v- 105Ile absent	.65 (.20 to 1.11), p=.005
Challenge x HNMT Thr105Ile 105Ile present–v- 105Ile absent	
Mix A –v- placebo	-.51 (-1.02 to -.001), p=.05
Mix B –v- placebo	-.45 (-.97 to .07), p=.091
Model 2 Intercept	-.63(-1.04 to -.22), p=.003
Challenge type	
Mix A –v- placebo	.39 (.12 to .65), p=.004
Mix B –v- placebo	.30 (.05 to .56), p=.020
Genotype - HNMT Thr105Ile 105Ile present –v- 105Ile absent	.51 (.08 to .95), p=.021
Challenge x HNMT Thr105Ile 105Ile present –v- 105Ile absent	
Mix A –v- placebo	-.53 (-1.04 to -.02), p=.041
Mix B –v- placebo	-.40 (-.92 to .12), p=.134
Week of study	
Wk 2 –v- Wk 6	.17 (-.05 to .40), p=.132
Wk 4 –v- Wk 6	.24 (.02 to .47), p=.035

Gender	.25 (-.04 to .53), p=.094
Baseline GHA score	.53 (.31 to .76), p<.001
Pre-trial diet	.06 (-.05 to .16), p=.301
Maternal education level	-.07 (-.37 to .22), p=.621
Maternal social class	-.21 (-.51 to .10), p=.184

Table 18.2. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (HNMT T939C) and Challenge type by Genotype interaction for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	≥85% consumption
Model 1	
Intercept	-.52 (-.78 to -.27), p<.001
Challenge type	
Mix A –v- placebo	.38 (.08 to .67), p=.013
Mix B –v- placebo	.29 (.002 to .57), p=.049
Genotype – HNMT T939C	
939C present –v- 939C absent	.55 (.12 to .97), p=.012
Challenge x HNMT T939C	
939C present –v- 939C absent	
Mix A –v- placebo	-.41 (-.89 to .07), p=.097
Mix B –v- placebo	-.26 (-.74 to .22), p=.290
Model 2	
Intercept	-.65 (-1.08 to -.21), p=.004
Challenge type	
Mix A –v- placebo	.42 (.13 to .71), p=.005
Mix B –v- placebo	.30 (.02 to .59), p=.036
Genotype – HNMT T939C	
939C present –v- 939C absent	.38 (-.03 to .79), p=.071
Challenge x HNMT T939C	
939C present –v- 939C absent	-.46 (-.94 to .02), p=.061
Mix A –v- placebo	-.23 (-.72 to .25), p=.338
Mix B –v- placebo	
Week of study	
Wk 2 –v- Wk 6	.18 (-.06 to .41), p=.140
Wk 4 –v- Wk 6	.26 (.03 to .49), p=.028
Gender	.22 (-.08 to .52), p=.150
Baseline GHA score	.54 (.31 to .77), p<.001

Pre-trial diet	.06 (-.05 to .17), p=.295
Maternal education level	-.06 (-.37 to .24), p=.677
Maternal social class	-.20 (-.52 to .11), p=.207

Table 18.3. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (COMT Val108Met) and Challenge type by Genotype interaction for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	≥85% consumption
Model 1	
Intercept	-.20 (-.61 to .22), p=.348
Challenge type	
Mix A –v- placebo	.42 (-.03 to .88), p=.067
Mix B –v- placebo	.11 (-.36 to .58), p=.645
Genotype – COMT Val108Met	
A present –v- A absent	-.16 (-.64 to .31), p=.493
Challenge x COMT Val108Met	
Val108 present –v- Val108 absent	
Mix A –v- placebo	-.30 (-.82 to .22), p=.261
Mix B –v- placebo	.09 (-.44 to .63), p=.730
Model 2	
Intercept	-.36 (-.87 to .16), p=.174
Challenge type	
Mix A –v- placebo	.39 (-.06 to .84), p=.087
Mix B –v- placebo	.11 (-.35 to .57), p=.645
Genotype – COMT Val108Met	
Val108 present –v- Val108 absent	-.17 (-.62 to .28), p=.458
Challenge x COMT Val108Met	
Val108 present –v- Val108 absent	
Mix A –v- placebo	-.23 (-.75 to .29), p=.382
Mix B –v- placebo	.12 (-.41 to .64), p=.662
Week of study	
Wk 2 –v- Wk 6	.15 (-.08 to .38), p=.196
Wk 4 –v- Wk 6	.20 (-.03 to .43), p=.083
Gender	.25 (-.04 to .54), p=.093

Baseline GHA score	.56 (.34 to .78), p<.001
Pre-trial diet	.05 (-.06 to .16), p=.360
Maternal education level	-.08 (-.38 to .21), p=.577
Maternal social class	-.17 (-.47 to .13), p=.268

Table 18.4. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (ADRDA2A C1921G) and Challenge type by Genotype interaction for 3 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	≥85% consumption
Model 1	
Intercept	-.16 (-.45 to .14), p=.293
Challenge type	
Mix A –v- placebo	.25 (-.08 to .57), p=.132
Mix B –v- placebo	.09 (-.24 to .42), p=.586
Genotype – ADRDA2A C1291G	
1291G present –v- 1291G absent	-.33 (-.73 to .08), p=.114
Challenge x ADRDA2A C1291G	
1291G present –v- 1291G absent	
Mix A –v- placebo	.01 (-.45 to .47), p=.962
Mix B –v- placebo	.19 (-.26 to .64), p=.413
Model 2	
Intercept	-.36 (-.81 to .08), p=.111
Challenge type	
Mix A –v- placebo	.27 (-.06 to .59), p=.107
Mix B –v- placebo	.10 (-.23 to .44), p=.540
Genotype - ADRDA2A C1291G	
G present –v- G absent	-.24 (-.62 to .14), p=.222
Challenge x ADRDA2A C1291G	
1291G present –v- 1291G absent	
Mix A –v- placebo	.01 (-.44 to .47), p=.959
Mix B –v- placebo	.20 (-.26 to .65), p=.389
Week of study	
Wk 2 –v- Wk 6	.12 (-.11 to .35), p=.306
Wk 4 –v- Wk 6	.22 (-.01 to .45), p=.057
Gender	.27 (-.02 to .56), p=.066
Baseline GHA score	.54 (.33 to .76), p<.001

Pre-trial diet	.06 (-.04 to .17), p=.239
Maternal education level	-.04 (-.34 to .25), p=.770
Maternal social class	-.24 (-.55 to .07), p=.122

19. Dietary infractions ('Whoops') analysis

A further analysis of interest was the effect of Challenge on GHA only for those children with no 'Whoops' recorded over the period of the study (n=113 children).

Table 19.1 shows a significant effect of Mix A on GHA. Juice was kept in a fridge and consumed at home either prior to the child's session in the EYS or on return from the EYS. Parents were therefore better able to keep a record of juice consumed. Similarly, the parent was usually aware of the food consumed at EYS which generally consisted of a portion of fruit and a drink of milk or water. 'Whoops' reports by parents therefore are likely to reflect dietary infractions which actually occurred over the challenge period for these very young children.

Table 19.1: MM Analysis of effect of Challenge, Baseline GHA and Pre-study Diet on GHA for 3yo children with ≥85% consumption and no Whoops

Mixed Model (n=120) -2LL= 633.81	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 134.85)=3.03 p=.052	Mix A	71	-.06 (1.16)	-.33 to .22	-.03 (.123)	-.27 to .21	A –v- P	.35	.07 to .64	.015
	Mix B	72	-.25 (1.04)	-.50 to -.01	-.19 (.123)	-.43 to .05	B –v- P	.19	-.10 to .48	.196
	Mix P	74	-.39 (1.18)	-.66 to -.12	-.38 (.121)	-.62 to -.14	-	-	-	-
Baseline GHA F(1, 117.74)=12.78 p=.001	-	-	-	-	-	-	-	.47	.21 to .73	.001
Pre-study diet F(1, 114.26)=.480 p=.490	-	-	-	-	-	-	-	.05	-.09 to .18	.490

20. Disaggregated measures analyses (≥85% consumption)

The GHA score represents an unweighted average of the standardised Parent, Teacher and Classroom Observation scores obtained in each of Mix A, Mix B and Placebo weeks. Further additional analyses of interest included the effect of Challenge on the disaggregated standardised GHA behaviour scores for each of the behaviour measures. Both Pre-study Diet and Baseline GHA were again included in these analyses. In each case, however, the Baseline GHA was that for the particular measure under analysis so that, for example, the parent Baseline GHA was included in the MM analysis of the effect of Challenge on Parent GHA. These analyses were carried out to help highlight the individual components of GHA over the challenge period which may have contributed to any relationship between consumption of additives and children's behaviour. However, it should be noted that each behaviour measure employed in the study focuses on slightly differing aspects of hyperactive behaviour over a range of differing contexts. The rationale for using an aggregated measure (GHA) to record behaviour in this study, in the first instance, was to obtain a behaviour score which reflected the child's behaviour over all of these differing situations. In addition, additive mixes were consumed at home and not at the EYS and any behaviour score within a particular context will be a function, amongst other things, not only of the amount of juice consumed but also time and individual differences in absorption of additives. Findings presented in Tables 20.1 to 20.3 below should be viewed in this light.

The analyses below show a significant effect of Mix A on Parent GHA but not for the Teacher GHA or Classroom Observation GHA. There is no significant effect of Mix B in any of the analyses although in the Parent GHA analysis this does approach significance (p=.085). The significance of these findings would require further investigation.

Table20.1: MM Analysis of effect of Challenge, Baseline Parent GHA and Pre-study Diet on Parent GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=130) -2LL= 1203.70	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge										
F(2, 211.38)=2.90	Mix A	102	.004 (1.77)	-.34 to .35	.02 (.168)	-.31 to .35	Mix A –v- P	.47	.07 to .87	.022
p=.057	Mix B	106	-.11 (1.71)	-.44 to .22	-.10 (.165)	-.43 to .22	Mix B –v- P	.35	-.05 to .74	.085
	Mix P	104	-.44 (1.67)	-.76 to -.11	-.45 (.166)	-.78 to -.12	-	-	-	-
Baseline Parent GHA F(1, 129.20)=.087 p=.769	-	-	-	-	-	-	-	-.05	-.36 to .27	.769
Pre-study diet F(1, 131.83)=.128 p=.721	-	-	-	-	-	-	-	-.03	-.21 to .15	.721

Table 20.2: MM Analysis of effect of Challenge, Baseline Teacher GHA and Pre-study Diet on Teacher GHA for 3yo children with ≥85% juice consumption

Mixed Model (n=131) -2LL = 576.39	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 189.42)=.568 p=.568	Mix A	101	-.27 (.86)	-.44 to -.10	-.22 (.063)	-.35 to -.10	Mix A –v- P	.02	-.11 to .15	.734
	Mix B	105	-.23 (.83)	-.39 to -.07	-.18 (.062)	-.30 to -.06	Mix B –v- P	.07	-.06 to .20	.297
	Mix P	106	-.24 (.86)	-.41 to -.08	-.25 (.062)	-.37 to -.12				
Baseline Teacher GHA F(1, 124.04)=134.47 p=.000	-	-	-	-	-	-	-	.60	.49 to .70	.000
Pre-study diet F(1, 118.99)=3.84 p=.211	-	-	-	-	-	-	-	.05	-.03 to .12	.211

Table 20.3: MM Analysis of effect of Challenge, Baseline COC GHA and Pre-study Diet on Classroom Observation GHA for 3yo children with ≥85% juice consumption

Mixed Model (131) -2LL=728.32	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 199.64)=1.06 p=.348			.							
	Mix A	102	.01 (.90)	-.16 to .19	.01 (.078)	-.15 to .16	Mix A –v- P	.08	-.09 to .25	.367
	Mix B	106	-.12 (.87)	-.28 to .05	-.12 (.077)	-.27 to .03	Mix B –v- P	-.05	-.22 to .12	.587
	Mix P	106	-.08 (.95)	-.26 to .11	-.07 (.077)	-.23 to .08				
Baseline COC GHA F(1, 123.86)=36.80 p=.000	-	-	-	-	-	-	-	.36	.24 to .48	.000
Pre-study diet F(1, 124.14)=6.31 p=.013	-	-	-	-	-	-	-	.11	.02 to .20	.013

21. Disaggregated measures analyses using all covariates

In order to provide parallel information to that for the effect of challenge on GHA the disaggregated measures wer examined again but now with all potential confounds controlled. The results are shown in Table 21.1 for all children, children with at least 85% juice consumption and a complete case analysis. The format of this table is slightly simplified to accommodate parameters for all the covariates.

Any single indicator is likely to be relatively less reliable compared to the aggregate measure. The consequent increased measurement error makes it less likely that a significant effect will be detected. For this reason the results are most appropriately discussed in terms of the effect sizes. For Challenge type under Model 2 16 of the 18 effect sizes in Tables 21.1 to 21.3 are in the direction of hyperactivity being more marked under the active than the placebo challenge. As before (Bateman et al, 2004) the strongest effects are found for parental ratings.

Table 21.1. General Linear Mixed Models estimates for **parent reported behaviour**[1] over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **3 year old** children (Model 1) and with the effects of week during study, gender, parent reported behaviour in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	Whole sample n=140	≥85% consumption n=130	Complete case n=73
Model 1			
Intercept	-.40 [-.69 to -.12] p =.006	-.45 [-.78 to -.13] p =.007	-.50 [-.91 to -.08] p =.020
Challenge type			
Mix A –v- placebo	.33 [-.00 to 0.67] p =.052	.47 [.07 to .86] p =.022	.55 [.06 to 1.03] p =.027
Mix B –v- placebo	.29 [-.05 to .62] p =.091	.35 [-.04 to .74] p =.081	.36 [-.12 to .84] p =.144
Model 2			
Intercept	-.19 [-.77 to .39] p =.526	-.28 [-.94 to .38] p =.409	-.18 [-1.05 to .70] p =.690
Challenge type			
Mix A –v- placebo	.33 [-.01 to .67] p =.058	.49 [.09 to .89] p =.016	.55 [.06 to 1.04] p =.027
Mix B –v- placebo	.27 [-.07 to .61] p =.117	.36 [-.04 to .76] p =.079	.37 [-.12 to .86] p =.138
Week of study			
Wk 2 –v- Wk 6	.05 [-.29 to .39] p =.775	.03 [-.37 to .44] p =.880	.15 [-.34 to .63] p =.556
Wk 4 –v- Wk 6	.16 [-.18 to .50] p =.359	.21 [-.19 to .61] p =.305	.09 [-.40 to .58 p =.724
Gender	-.08 [-.51 to .36] p =.724	.09 [-.39 to .56] p =.722	-.25 [-.90 to .40] p =.442
Baseline GHA score	-.04 [-.30 to .21] p =.737	-.03 [-.35 to .29] p =.840	-.08 [-.55 to .39] p =.737
Pre-trial diet	-.03 [-.20 to .13] p =.686	-.01 [-.19 to .17] p =.910	-.01 [-.26 to .24] p =.931
Maternal education level	-.11 [-.58 to .35] p =.628	-.20 [-.71 to .31] p =.442	-.11 [-.81 to .60] p =.764
Maternal social class	-.20 [-.68 to .27] p =.397	-.33 [-.85 to .19] p =.206	-.35 [-1.03 to .33] p =.312

Table 21.2. General Linear Mixed Models estimates for **teacher reported behaviour**[2] over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **3 year old** children (Model 1) and with the effects of week during study, gender, teacher reported behaviour in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	Whole sample n=140	≥85% consumption n=130	Complete case n=73
Model 1			
Intercept	-.17 [-.33 to -.02] p =.027	-.21 [-.36 to -.05] p =.012	-.43 [-.60 to -.26] p <.001
Challenge type			
Mix A –v- placebo	.004 (-.11 to .12] p =.944	.03 [-.11 to .16] p =.701	.08 [-.06 to .22] p =.264
Mix B –v- placebo	.05 [-.08 to .16] p =.448	.07 [-.06 to .20] p =.317	.10 [-.04 to .25] p =.157
Model 2			
Intercept	-.38 [-.62 to -.13] p =.003	-.37 [-.63 to -.10] p =.007	-.44 [-.74 to -.14] p =.005
Challenge type			
Mix A –v- placebo	.01 [-.11 to .13] p =.861	.03 [-.11 to .16] p =.688	.09 [-.06 to .23] p =.245
Mix B –v- placebo	.06 [-.06 to .17] p =.338	.08 [-.05 to .21] p =.230	.10 [-.04 to .25] p =.167
Week of study			
Wk 2 –v- Wk 6	.03 [-.08 to .15] p =.572	.00 [-.13 to .14] p =.988	-.04 [-.18 to .11] p =.631
Wk 4 –v- Wk 6	-.001 [-.12 to .12] p =.993	-.02 [-.15 to .11] p =.764	.02 [-.12 to .17] p =.769
Gender	.14 [-.05 to .34] p =.148	.15 [-.05 to .35] p =.147	.05 [-.17 to .27] p =. 659
Baseline GHA score	.57 [.47 to .67] p <.001	.58 [.47 to .68] p <.001	.50 [.37 to .63] p <.001
Pre-trial diet	.06 [-.02 to .13] p =.125	.04 [-.04 to .11] p =.330	.05 [-.04 to .13] p =.291
Maternal education level	.14 [-.07 to .34] p =.180	.14 [-.07 to .35] p =.202	.04 [-.20 to .28] p =.753
Maternal social class	-.04 [-.25 to .16] p =.685	-.01 [-.23 to .20] p =.908	-.02 [-.25 to .22] p =.893

Table 21.3. General Linear Mixed Models estimates for **classroom observation**[3] over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **3 year old** children (Model 1) and with the effects of week during study, gender, classroom observation in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	Whole sample n=140	≥85% consumption n=130	Complete case n=73
Model 1			
Intercept	-.08 [-.23 to .07] p =.314	-.06 [-.23 to .11] p =.497	-.04 [-.25 to .16] p =.669
Challenge type			
Mix A –v- placebo	.09 [-.06 to .23] p =.231	.07 [-.11 to .24] p =.457	.05 [-.15 to .26] p =.598
Mix B –v- placebo	-.02 [-.16 to .13] p =.802	-.06 [-.23 to .12] p =.525	-.05 [-.25 to .15] p =.600
Model 2			
Intercept	-.46 [-.74 to -.17] p =.002	-.51 [-.83 to -.18] p =.002	-.53 [-.92 to -.14] p =.008
Challenge type			
Mix A –v- placebo	.09 [-.06 to .23] p =.240	.10 [-.07 to .27] p =.262	.08 [-.11 to .28] p =.420
Mix B –v- placebo	.001 [-.14 to .14] p =.988	-.01 [-.18 to .16] p =.924	-.02 [-.21 to .18] p =.865
Week of study			
Wk 2 –v- Wk 6	.22 [.08 to .36] p =.002	.32 [.15 to .49] p <.001	.33 [.14 to .53] p =.001
Wk 4 –v- Wk 6	.19 [.04 to .33] p =.011	.29 [.12 to .46] p =.001	.30 [.10 to .49] p =.003
Gender	.22 [-.00 to .45] p =.054	.19 [-.06 to .43] p =.130	.22 [-.07 to .51] p =.136
Baseline GHA score	.36 [.24 to .47] p <.001	.34 [.22 to .47] p <.001	.34 [.19 to .49] p <.001
Pre-trial diet	.13 [.04 to .21] p =.003	.12 [.03 to .22] p =.008	.15 [.04 to .26] p =.009
Maternal education level	-.07 [-.31 to .16] p =.539	-.08 [-.33 to .17] p =.511	.01 [-.31 to .32] p =.969
Maternal social class	-.07 [-.30 to .17] p =.558	-.06 [-.31 to .20] p =.653	-.12 [-.42 to .18] p =.439

22. Responders and non-responders to additives analyses

An additional question of interest relates to responders and non-responders to the additives and whether the observed differences in responses to Mix A. For those children who completed the study, the difference scores between Mix A and Mix P (mean .14, SD 1.72 [IQR -.53 to .99]) and those for Mix B and Mix P (mean .06, SD 1.33 [IQR .04 to .79]) .were calculated. Children with scores $\geq 75^{th}$ percentile were classed as 'responders', those $\leq 25^{th}$ percentile as 'non-responders', and the remaining children as 'neutral' responders. Of the 34 children who were responders to Mix A and 34 children who were responders to Mix B, 19/49 (38.8%) of responders to Mix A and/or Mix B were responders to both mixes, 15/49 (30.6%) were responders to Mix A only and 15/49 (30.6%) were responders to Mix B only.

23. References

Abikoff H, Gittelman R. (1985). Classroom observation code: A modification of the Stony Brook code. *Psychopharmacology Bulletin, 21,* 901-909.

Bateman B, Warner JO, Hutchinson E, Dean T, Rowlandson P, Gant C, Grundy J, Fitzgerald C, Stevenson J. (2004). The effects of a double blind, placebo controlled, artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. *Archives of Disease in Childhood, 89*, 506-11.

Conners CK, Epstein JN, Angold A, Klaric J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology, 31(5),* 555-562.

Conners CK. (1994). *The Conners continuous performance test.* Toronto, Canada: Multi-Health Systems.

Conners CK (1997. *Conners Rating Scales Revised: Technical Manual*. Toronto, Canada: Multi-Health Systems

Cook RJ, Zeng L, Yi GY. (2004). Marginal analysis of incomplete longitudinal binary data: A cautionary note on LOCF imputation. *Biometrics, 60,* 820-828.

DuPaul GJ, Anastopoulos AD, Power TJ, Murphy K, Barkley RA. (1994). *ADHD Rating Scale-IV,* Unpublished rating scale.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey K, Ikeda M. (1997). Teacher ratings of ADHD symptoms: Factor structure and normative data. *Psychological Assessment, 9,* 436-444.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. (1998). *AD/HD Rating Scale IV: Checklists, Norms and Clinical Interpretation*. New York; Guilford Press.

Epstein N, Erkanli A, Conners CK, Klaric J, Costello JE, Angold A. (2003) Relations between continuous performance test performance meaures and ADHD behaviours. *Journal of Abnormal Child Psychology, 31(5), 543-554.*

Gueorguieva R, Krystal JH. (2004). Move over ANOVA: Progress in analysing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry, 61,* 310-317.

Mallinckrodt CH, Sanger TM, Dubé S, DeBrota DJ, Molenberghs G, Carroll RJ, Potter WZ, Tollefson GD. (2003). Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. *Biological Psychiatry, 53,* 754-760.

Mallinckrodt et al. (2004). Choice of the primary analysis in longitudinal clinical trials. *Pharmaceutical Statistics, 3,* 161-169.

Molenberghs et al. (2004). Analysing incomplete longitudinal clinical trial data (with discussion). *Biostatistics, 5,* 445-464.

Annex 2

Food and Behaviour in Children Study

8-9yo Report to FSA

This set of material describes in detail the design, methods, data analysis and findings of the study of 8-9 year olds. It is accompanied by a parallel report for the study of 3 year old children.

Annex 2 CONTENTS

1. Foreword

The first part of this report presents data on the enlistment of schools (Table 2.1), parents and children (Figure 3.1). Data on schools, school roll, the number of Year 3 and 4 classes and free school meal uptake (fsm%) indicates that schools in our sample are representative of all schools within the Southampton City area. Information relating to characteristics of the whole sample and stratified by gender are presented in Table 3.1. In Table 3.2 we also present characteristics for a subgroup of children with ≥85% juice consumption in any challenge week and separately for a similar but smaller group (complete case) with ≥85% consumption of juice and no missing behaviour data. This first part of the report, therefore, also provides data on the percentage consumption of juices over the period of the study (Section 5). These subgroups have been employed in the Mixed Model analyses from Section 10 onwards where more detailed comments regarding analyses can be found. Details of the genotype analysis to be carried out are provided in Section 4. Summary tables are also provided relating to the child's diet i.e. dietary intake prior to the start of the study. This was based on maternal recall of the child's diet in the previous 24 hours (Section 6) and then classified in terms of the number of foods containing the food additives under study. This factor was also taken into account in analyses. Further tables provide data relating to the incidence of dietary infractions ('Whoops') reported by parent in a daily diary which was collected weekly (Section 19). This data was employed in secondary analyses, one of which focused only on those children with ≥85% consumption of juice and no recorded 'Whoops'.

Information is provided on the measures of behaviour employed in this 8-9yo study (Section 7) together with summary tables relating to the distribution of behaviour data collected prior to, at baseline and over the period of the study (Section 9.3). Behaviour screen data collected from teachers in participating schools prior to the start of the study also indicated that behaviour ratings for our sample are representative of those for children of the same age in participating schools.

There is also comment on the use of Mixed Model methods in data analysis. A discussion of the use of an appropriate covariance matrix structure in the Mixed Models approach follows in Section 10 before discussion and presentation of the findings of the main Mixed Model analyses and analyses based on Genotype (Section 11 onwards). Additional secondary analyses of interest focussing on disaggregated behaviour measures and responders and non-responders to additives can be found in Section 20 and 21 and 22. Section 23 reports on the Acute Challenge phase of the study.

2. Schools

In this 8-9yo study information on free school meals (fsm%) was obtained for all schools in the Southampton City Council area. Schools were then ranked into five groups (Group 1: 0-10%, Group 2: 11-20%; Group 3: 21-30%; Group 4: 31-40%; Group 5: 41 to 50%). A fixed proportion of schools within each group were approached and enlisted in order to ensure a study sample reflecting the full range of socio-economic background of children in the area (Table 2.1).

Table 2.1 Characteristics of participating schools

School group based on free school meal uptake (fsm%) in 2004	Soton schools 2005 n (%)	fsm uptake 2004 %	Projected Y4 roll 2005 n (%)	Target enlist-ment schools n (%)	Schools enlisted n (%)	fsm uptake %	Target enlistment pupils n=135	Pupils enlisted n=144(%)
Group 1 (0-10%)	6 (15.4)	6.7	434 (18.3)	1.4 (15.4)	1 (11.1)	6	24.7	20 (13.9)
Group 2 (11-20%)	11 (28.2)	15.7	875 (36.8)	2.5 (28.2)	3 (33.3)	12	49.7	66 (45.8)
Group 3 (21-30%)	9 (23.1)	26.6	561 (23.6)	2.1 (23.1)	2 (22.2)	25	31.9	36 (25.0)
Group 4 (31-40%)	4 (10.3)	33.8	140 (5.9)	.9 (10.3)	1 (11.1)	31	8	5 (3.5)
Group 5 (41-50%)	9 (23.1)	43.3	365 (15.4)	2.1 (23.1)	2 (22.2)	46	20.7	17 (11.8)
Total	39	25	2375	9	9	24	135	144

Of 39 primary/junior schools in the Southampton City area, 11 were approached and 9 agreed to participate. After obtaining parental consent, children of the target age were screened using the behaviour rating scales. Table 2.2 shows dates when the Research Team were present in individual schools to carry out the study.

Table 2.2: Periods in schools

Period 1: Mon 11th April to Fri 27th May 2005 Fairisle Jnr (Gp3 –fsm 28%); Newlands Pr (Gp5 – fsm 49%); St Denys Pr (Gp 4 – fsm 31%); St. Mary's ((Gp5 – fsm 44%)
Period 2: Mon 23rd May to Fri 22nd July 2005; Bitterne Park Jnr (Gp2 – fsm 13%); Ludlow Jnr (Gp2 – fsm 12%)
Period 3: Mon 5th September to Fri 21st October 2005 Springhill Catholic Pr (Gp2 – fsm 11%); St Monica Jr (Gp1 – fsm 6%); Tanners Brook Jnr (Gp3 – fsm 22%)

3. Participants

Behaviour screen questionnaires were completed for children of the target age by teachers in participating schools. Parents were subsequently approached via the school to participate in the study. Those who returned an expression of interest form were contacted by phone and a convenient time for a home visit arranged. On this visit, a research assistant and the study dietitian provided full information about the study and its dietary implications and informed consent was obtained. At the home visit the study dietitian also obtained a report based on 24-hour recall by the parent of the child's pre-study diet which allowed an assessment of baseline levels of the number of foods containing additives consumed by the child in the previous 24 hour period. Figure 3.1 provides details of enlistment of parents and children.

At this stage, 5 children were excluded from the study because of the presence of allergic reactions to food (n=2) or blackcurrant juice (n=2) and in the case of 1 child who would be going on holiday in the challenge period. A further 9 children were not enlisted to the study or refused to participate prior to or at the home visit (no reason given [n=4], moving home [n=2], no contact obtained [n=4], consent subsequently refused by father [n=1, concern about diet change [n=1], concern about child eating additives/drinking juice [n=1]). Of the 146 parents who consented to participate at the home visit, 2 (1.4%) subsequently refused for reasons related to juice consumption. The children who were successfully enlisted (n=144) had a mean age of 106.3m (SD 5.9m, range 93 to 123m) with a similar age distribution being found for 69 girls (mean age 106.1m, SD 5.8m, range 96 to 116m) and 75 boys (mean 106.4m, SD 6.1m, range 93 to 123m). Table 3.1 presents detail relating to characteristics of all participating parents. For data analysis purposes, Table 3.2 also presents the same information for the subgroup of children (n=119) who had consumed $\geq$85% juice in any challenge week over the period of the study and for a smaller group of these children (n=91) who also had no missing behaviour data. Of the 144 participating children, 14 (9.7%) failed to complete the study for reasons related to juice consumption (n=9), parental pressure of work or other commitments (n=2) and medical reasons (n=3). No differences were found in terms of age or gender between those children who completed the study and the group of children who did not complete (9 boys, 7 girls) with the exception that there were a lesser proportion of widowed, divorced or single parents in the latter group (p=.01).

Figure 3.1: Enlistment of parents and 8-9YO children to main study

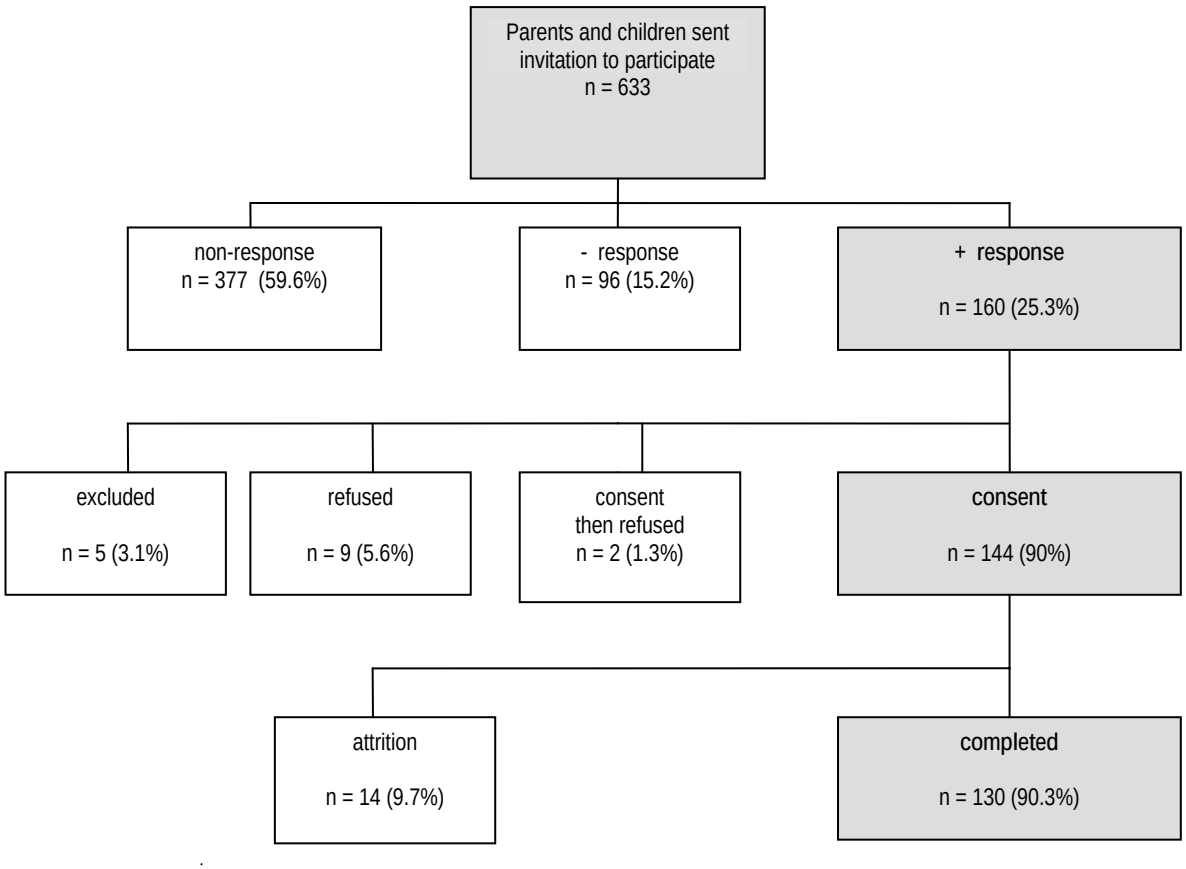


Table 3.1: Characteristics of sample

	Total N=144	Boys n=75 n(%)	Girls n=69 n(%)	Diffs $\chi 2$(df)p=
Age (children)				
≤ 9yrs	90 (62.5)	47 (62.7)	43 (62.3)	ns
>9yrs	54 (37.5)	28 (37.3)	26 (37.7)	
Racial background (parents)				
White	130 (90.3)	69 (92.0)	61 (88.4)	ns
Other	14 (9.7)	6 (8.0)	8 (11.6)	
Marital status				
Married/partner	115 (79.9)	61 (81.3)	54 (78.3)	ns
Single/separated/divorced/widowed	29 (20.1)	14 (18.7)	15 (21.7)	
National Statistics Social Class (Father)				
Higher occupations	37 (25.7)	17 (22.6)	20 (29.0)	
Intermediate occupations	20 (13.9)	11 (14.6)	9 (13.0)	
Lower occupations	45 (31.3)	25 (33.3)	20 (29.0)	ns
Never worked/long term unemployed	9 (6.3)	7 (9.3)	2 (2.9)	
Missing data	33 (22.9)	15 (20.0)	18 (26.1)	
National Statistics Social Class (Mother)				
Higher occupations	38 (26.4)	21 (28.0)	17 (24.6)	
Intermediate occupations	26 (18.1)	12 (16.0)	14 (20.3)	
Lower occupations	32 (22.2)	17 (22.7)	15 (21.7)	ns
Never worked/long term unemployed	32 (22.2)	19 (25.3)	13 (18.8)	
Missing data	16 (11.1)	6 (8.0)	10 (14.5)	
Educational qualifications (Mother)				
School to 16 (no qualifications /				
Certificates below ‘A’ level)	60	32 (42.7)	28 (40.6)	
‘A’ levels	42	21 (28.0)	21 (30.4)	ns
University Degree/Post-graduate qualification	27	16 (21.3)	11 (15.9)	
Missing data	15	6 (8.0)	9 (13.0)	

Table 3.2: Characteristics of sample (all children, ≥85% consumption, complete case)

	Children in total sample analysis n =144 n(%)	Children in ≥85% drunk analysis n = 119 n(%)	Children in complete case analysis n = 91 n(%)
Racial background			
White	130 (90.3)	110 (92.4)	85 (93.4)
Other	14 (9.7)	9 (7.6)	6 (6.6)
Missing data	-	-	-
Marital status			
Married/partner	115 (79.9)	99 (83.2)	80 (87.9)
Single/separated/divorced/widowed	29 (20.1)	20 (16.8)	11 (12.1)
NSSC* (Father)			
Higher occupations	37 (25.7)	35 (29.4)	29 (31.9)
Intermediate occupations	18 (12.5)	17 (14.3)	14 (15.4)
Lower occupations	44 (30.6)	39 (32.8)	30 (33.0)
Never worked/long term unemployed	7 (4.9)	7 (5.9)	6 (6.6)
No father present	29 (20.1)	20 (16.8)	11 (12.1)
Missing data	9 (6.3)	1 (0.8)	1 (1.1)
NSSC(Mother)			
Higher occupations	38 (26.4)	35 (29.4)	27 (29.7)
Intermediate occupations	26 (18.1)	25 (21.0)	20 (22.0)
Lower occupations	32 (22.2)	29 (24.4)	25 (27.5)
Never worked/long term unemployed	32 (22.2)	28 (23.5)	19 (20.9)
Missing data	16 (11.1)	2 (1.7)	-
Mother's education			
School to 16 (no qualifications / certificates below 'A' level)	60 (41.7)	54 (45.3)	40 (44.0)
'A' levels	42 (29.2)	39 (32.8)	32 (35.2)
University Degree/Post-graduate qualification	27 (18.8)	24 (20.2)	19 (20.9)
Missing data	15 (10.4)	2 (1.7)	-

*NSSC: National Statistics Social Class

4. Genotyping

In order to help identify which processes mediate the effect of AFCs on hyperactivity, the present study also examines whether the effect of artificial food colors (AFCs) on hyperactivity is moderated by genetic difference between children.

Genetic polymorphisms were selected from the dopamine (catechol-o-methyltransferase) and adrenergic (adrenergic receptor alpha 2A) neurotransmitter systems since these have previously been implicated in ADHD. Since there is also a suggestion that histamine may be involved in the effects of AFCs, genetic polymorphisms from this system were also included.

Given the difficulties in genotyping variable number tandem repeat (VNTR) polymorphisms using DNA from cheek cells (the only means of access to DNA for general population samples of children) polymorphisms were selected for the present analysis that were single nucleotide polymorphisms (SNPs). Consequently results are presented here for two SNPs in the histamine N-methyltranferase gene (HNMT Thr105Ile) and HNMT T939C), for one SNP in the dopamine gene, catechol-o-methyltransferase (COMT Val108Met), and one SNP in the adrenergic neurotransmitter system adrenergic receptor, alpha 2A (ADRA2A C1292G).

Table 4.1 presents the distribution of alleles present and absent for each of the genetic polymorphisms outlined both for all children and by gender.

Table 4.1: Genotyping by Gender

	COMT Val108Met %	HNMT T939C %	ADRA2A C1291G %	HNMT Thr105Ile %
Total Sample n = 144	Val108 present = 62.5 Val108 absent = 28.5 Missing = 9.0	939C present = 37.5 939C absent = 53.5 Missing = 9.0	1291G present = 40.3 1291G absent = 48.6 Missing = 11.1	105Ile present = 19.4 105Ile absent = 71.5 Missing = 9.0
Boys (n=75)	Val108 present = 57.3 Val108 absent = 34.7 Missing = 8.0	939C present = 37.3 939C absent = 56.0 Missing = 6.7	1291G present = 32.0 1291G absent = 60.0 Missing = 8.0	105Ile present = 17.3 105Ile absent = 76.0 Missing = 6.7
Girls (n=69)	Val108 present = 68.1 Val108 absent = 21.7 Missing = 10.1	939C present = 37.7 939C absent = 50.7 Missing = 11.6	1291G present = 49.3 1291G absent = 36.2 Missing = 14.5	105Ile present = 21.7 105Ile absent = 66.7 Missing = 11.6

5. Consumption of juice

Parents completed a daily diary of juice consumption and compliance with the diet over the study period. At the end of each week of the study, any bottles containing juice not consumed in the previous week were collected, returned to the study office and the contents measured to help validate, where possible, parental reports of juice consumption by children. Consumption of juice remained at an acceptable level over the period of the study for the majority of children. Of the children who completed the study, 98/130 (75%) consumed more than 85% of juices (at least 6 out of 7 daily drinks per week) and a further 13/130 (10%) consumed more than two-thirds of all drinks. Table 5.1 presents detail of juice consumption over weeks of the study and by challenge.

Table 5.1: Consumption of juice by Week and Challenge

	Week 2 n (%)	Week 4	Week 6	Mix A n (%)	Mix B	Placebo
≥ 85 %	107 (74.3)	109 (75.7)	103 (71.5)	104 (72.2)	112 (77.8)	103 (71.5)
60- 84 %	13 (9.0)	8 (5.6)	12 (8.3)	16 (11.1)	8 (5.6)	9 (6.3)
< 60 %	9 (6.3)	13 (9.0)	14 (9.7)	11 (7.6)	11 (7.6)	15 (10.4)
Missing	15 (10.4)	14 (9.7)	15 (10.4)	10 (6.9)	13 (9.0)	17 (11.8)
				14 (9.7)		

6. Dietary Measures

Pre-study dietary measures were collected in order to investigate if diet prior to the study moderated the effect of challenge. Pre study dietary measures are based on 24 hour parental recall of consumption of food items containing additives prior to the start of the study. Table 6.1 presents data relating to the distribution of the number of food items containing additives consumed by all children and by gender groups.

6.1 Dietary intake (24hr recall) of foods containing additives by Gender and Age group

No. of foods	Total Sample n (%)	Boys n (%)	Girls N (%)
0	13 (9.0)	6 (8.0)	7 (10.1)
1	30 (20.8)	14 (18.7)	16 (23.2)
2	25 (17.4)	15 (20.0)	10 (14.5)
3	50 (34.7)	25 (33.3)	25 (36.2)
4	17 (11.8)	9 (12.0)	8 (11.6)
5	6 (4.2)	5 (6.7)	1 (1.4)
6	3 (2.1)	1 (1.3)	2 (2.9)

Dietary infractions ('Whoops')

In the weekly diary, completed daily, parents also recorded a 'Whoops' event each time in any one day the child consumed a portion of food containing artificial colours or sodium benzoate. The occurrence of dietary infractions or 'mistakes' by children over the study period was low (25% = 0; 41% = 1 to 2; 16.9% = 3 to 4; 17% > 4). Rates did not differ during active and placebo weeks (Table 6.2).

Table 6.2 Dietary infractions ('Whoops') by week and challenge

	Over Weeks			Over Challenge		
	whoops in week 2 n(%)	whoops in week 4 n(%)	whoops in week 6 n(%)	whoops in mix A weeks n(%)	whoops in mix B weeks n(%)	whoops in placebo weeks n(%)
TOTAL						
0	95 (66)	97 (67)	87 (60)	94 (65)	90 (63)	95 (66)
1	17 (12)	22 (15)	28 (19)	25 (17)	24 (17)	18 (13)
>1	13 (9)	8 (6)	7 (5)	7 (5)	13 (9)	8 (6)
Missing	19 (13)	17 (12)	22 (15)	18 (13)	17 (12)	23 (16)
BOYS						
0	56 (75)	53 (71)	49 (65)	52 (69)	53 (71)	53 (71)
1	7 (9)	10 (13)	14 (19)	10 (13)	11 (15)	10 (13)
>1	3 (4)	3 (4)	1 (1)	3 (4)	3 (4)	1 (1)
Missing	9 (12)	9 (12)	11 (15)	10 (13)	8 (11)	11 (15)
GIRLS						
0	39 (57)	44 (64)	38 (55)	42 (61)	37 (54)	42 (61)
1	10 (15)	12 (17)	14 (20)	15 (22)	13 (19)	8 (12)
>1	10 (15)	5 (7)	6 (9)	4 (6)	10 (15)	7 (10)
Missing	10 (15)	8 (12)	11 (16)	8 (12)	9 (13)	12 (17)

7. Behaviour Measures

Behaviour screen measures (Teacher and Parent)

ADHD Rating Scale – IV (Home version) (DuPaul, Anastopoulos ,Power et al, 1994)
This measure provides a total score for 18 items (inattention n=9, hyperactive n=9) directly adapted from the ADHD symptom list as specified in the DSM-IV and scored on a scale of 0 (never or rarely) to 3 (very often) so that parent may best describe the frequency of occurrence of each specified behaviour over the past 6 months. This questionnaire completed by all participating parents prior to baseline (Week 0) is similar in format to and correlates significantly with the ADHD Rating Scale – IV (Teacher version). The parent scale has been shown to have acceptable psychometric properties including inter-rater reliability, test-retest reliability and internal consistency (Zhang, Aries Vowles, Michelson, 2005; DuPaul, Power, Anastopoulos & Reid, 1998). Normative data has been published (DuPaul, Anastopoulos, Power et al, 1998). Scores also have adequate positive and negative predictive power in the diagnosis of ADHD (Power, Doherty, Panichelli-Mindel et al, 1998).

ADHD Rating Scale – IV (Teacher version) (DuPaul, Power, Anastopoulos et al, 1997)
This 18-item questionnaire is similar in format to the parent/home version. This behaviour screen questionnaire was completed by teachers of all children of the target age within participating schools. Normative data has been published (DuPaul, Power, Anastopoulos et al, 1997). As with the Parent/Home version, the ADHD Rating Scale-IV manual presents information on normative data and psychometric properties of this scale.

Behaviour weekly measures (Teacher and Parent)

ADHD Rating Scale – IV (Teacher version) (DuPaul, Power, Anastopoulos et al, 1997)
This questionnaire (see Behaviour Screening above) was completed by the teacher to describe the frequency of the specific behaviour displayed *over the past week* for each week of the study (Week 0 [baseline] to Week 6). Only 10 of the 18 items (inattention n=5, hyperactivity-impulsivity n=5) were completed on a weekly basis and a total score obtained.

ADHD Rating Scale – IV (Home/Parent version) (DuPaul, Anastopoulos, Power et al, 1994)
Similar in format to the weekly teacher behaviour questionnaire, this 10-item questionnaire was completed by all participating parents each week of the study (Week 0 [baseline] to Week 6).

Classroom observation

Classroom Observation Code (COC) (Abikoff & Gittelman, 1985)

The COC is one of the earliest and most thoroughly evaluated school observation coding systems. The COC assesses the occurrence of 12 mutually exclusive behaviours during structured didactic teaching and during periods of independent work under teacher supervision. A full description of the behaviours coded and scored is given in the COC training manual and these include: interference (to others); interference to teacher; off-task; non-compliance; aggression (physical); verbal aggression to teacher; minor motor movement; gross motor – vigorous; out-of-chair and solicitation of teacher. Each child was observed for a total duration of 24 minutes each week (3 observation sessions x 8 minutes duration) and a total weekly score was derived from the total score over each session. The COC has adequate interobserver reliability, discriminates between hyperactive and nonhyperactive children and has no detectable observer effect on child behaviour (Abikoff, Gittelman-Klein, Klein, 1977; Abikoff, Gittelman & Klein, 1980).)

Response inhibition and attention

Conners' Continuous Performance Test II (CPTII) (Conners, 1994)

The CPTII is a visual paradigm of 14 minutes duration and is used to evaluate attention and the response inhibition component of executive control. The CPT can be administered to children 6years of age and older. The participant is presented with 360 trials in 18 blocks each of 20 trials. The blocks differ in terms of inter-stimulus interval (ISI) condition (1, 2 or 4 seconds) so that the CPT task is presented in 6 consecutive time blocks with each time block containing all 3 randomly presented ISI conditions. The CPT generates multiple dependent measures including hit reaction time (RT), standard error of reaction time (RT [SE]), % of omission errors, % of commission errors, and the signal detection parameters of signal dectability or d-prime (*d'*) and response bias or beta (β). Unlike traditional CPTs, this CPT-not X task, requires the participant to press a computer key immediately in response to all letter presentations other than the letter 'X' thus making the target the more frequently occurring signal. More frequent responding and the availability of a greater number of trials therefore tends to elicit more errors of commission, places a greater demand on response inhibition than more conventional CPT task , allows for more accurate data on RT and RT (SE) and thus provides more reliable measures relating to attentional function and detectability. For the purposes of this study, only 4 of the 6 measures (standard error of reaction time (RT [SE]), % of commission errors, *d'* and β.were used to derive a weekly CPT aggregate score. In a recent study which published normative data on the CPT (Epstein N, Erkanli A, Conners CK et al 2003)., these measures have been shown to be highly correlated with the ADHD Rating Scale parental measure of behaviour employed in this study. Further information relating to normative data and reliability estimates are provided in the CPTII manual (Conners, 2000).

8. Data analysis methods

Generating the behaviour Global Hyperactivity Aggregate (GHA)

The weekly scores from the Teacher, Parent and COC measures for each child were standardised to Time 0 at baseline [T0]) for the same measure –

$$\text{Weekly standardised (z) aggregate score} = \frac{\text{(score X – mean X at T0)}}{\text{SD at T0}}$$

In the case of the weekly aggregate CPT score, this is obtained by first standardising the score for each of the 4 scales to T0, aggregating the scores and restandardising to T0 z-score to obtain a weekly CPT z-score. The primary outcome measure, the global hyperactivity aggregate (GHA), is a similarly unweighted aggregate of the weekly Teacher, Parent, COC and CPT z-scores. The GHA is calculated only when at least 3 of these behaviour scores are present for any week.

Use of Mixed Models method

The effect of challenge (Mix A –v- Placebo; Mix B –v- Placebo) on GHA was tested using Mixed Models analyses. A number of reports now suggest that traditional methods of analysing repeated measures data, such as end-point analysis and univariate/multivariate repeated measures analysis of variance, have a number of disadvantages (Gueorguieva & Krystal, 2004; Mallinckrodt, Watkin Molenberghs & Carrol, 2004). Mixed effects models offer a number of advantages via greater flexibility and use of all the available data. Such approaches are able to handle unequally spaced observations over time, correlation between repeated measures on the same subject over time through a random effects approach, differing group sizes and missing data, as long as such data is missing at random. The use of a mixed models analysis also allows a random-effects approach to be combined with patterns of variance in the data over time through specification of a covariance pattern model. For example, selection of a *compound symmetry* covariance structure involves estimation of a small number of parameters by assuming both equal variances at all time points and equal correlation between any two measurements on the same subject. Specification of an *autoregressive* structure of intermediate complexity assumes decreasing correlation between pairs of measures with increasing time difference. An *unstructured* covariance model sets no restrictions at all on the covariance structure but involves estimation of a larger number of parameters.

The present study employed a mixed model analysis to examine the effect of challenge on GHA since missing data was viewed as missing at random and not related to any observed change of behaviour in the child (see Section 2. Participants). Likelihood ratio tests, reduced SE and covariance parameters indicated that an *unstructured* model was the best fit model to apply to the data in the present study.

Power calculation

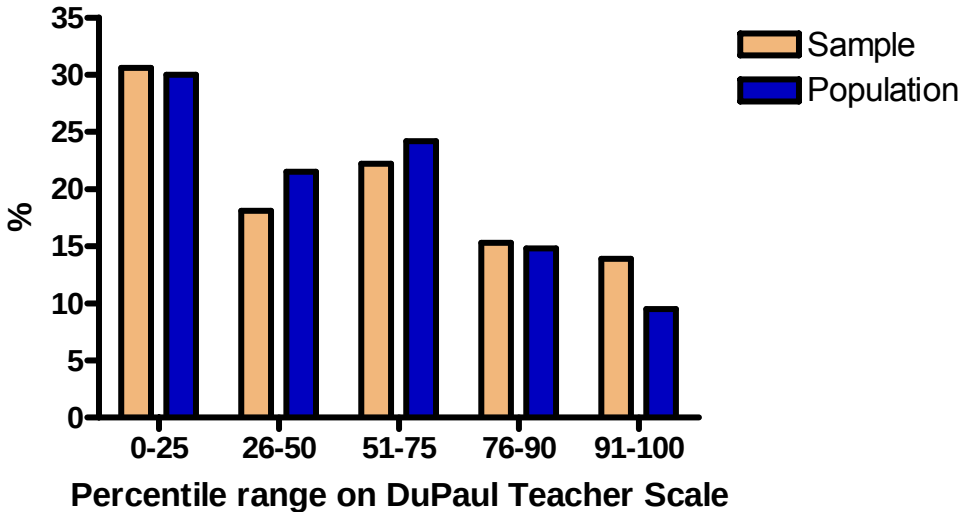
The study was powered to detect differences between the active and placebo periods and accordingly in each case the effects of Mix A and Mix B were compared to the effect of placebo. It was anticipated that the additional controls on placebo effects would result in an effect size that would be smaller than that achieved in the Isle of Wight study (Bateman et al., 2004). With a sample of 80 children there was 80% power at alpha = .05 to identify an effect size of 0.32 i.e. the magnitude of the difference in GHA mean score changes (in standard deviation units). The latter being somewhat lower than that achieved in the previous study (0.51). There were uncertainties over the number of children and families who would comply with the demands of a 7 week study and so a target of 120 children was set to reduce the impact of attrition on power and this target was eventually exceeded in both age groups.

9. Behaviour Data

Behaviour at screen

Figure 9.1 Presents distribution of scores over percentile ranges for children in the population of 8-9 year olds in participating schools and in the FABiC sample.

Figure 9.1: Teacher ratings for all 8/9yo children in participating schools



There was no significant difference between the proportions of children in each the percentile ranges in the FABiC sample and in the rest of the School population (χ^2 (4) = 5.05, p <.283). The FABiC sample therefore had achieved the approximately equal distribution across these percentile ranges aimed for and were representative of the general population in terms of the Teacher score. The percentile ranges for the raw DuPaul Teacher Scale scores were calculated separately for males and females. This resulted in an approximately equal number of male (n = 75) and females (n = 69) in the FABiC sample. There was no significant difference between the proportion of males and females in the FABiC sample and in the rest of the School population (χ^2 [1] = 0.59, p <.445).

Table 9.1 provides data on behaviour scores for the sample and population by gender for the sample and population in participating schools.

Table 9.1: Teacher scores for sample –v- population of 8-9 yo children in participating schools

	Teacher Behaviour Scores					
	Boys		Girls		All	
	Sample (n=75) Median [IQR]	Population (n=235) Median [IQR]	Sample (n=69) Median [IQR]	Population (n=250) Median [IQR]	Sample (n=144) Median [IQR]	Population (n=485) Median [IQR]
Hyperactivit	3	3	1	0	2	1

y	[0 to 13]	[0 to 8]	0 to 3	0 to 2	0 to 7.75	0 to 5
	p=.441, ns		P=.302, ns		p=.223, ns	
Inattention	7 [1 to 16]	7 [2 to 12]	2 0 to 6	1 0 to 7	4 0 to 11.75	4 0 to 9.5
	p=.883, ns		P=.731, ns		p=.789, ns	
Total behaviour	11 [1 to 30]	10 [2 to 20]	3 0 to 9.5	2 0 to 9	5 1 to 19.75	6 1 to 15
	p=.767, ns		P=.453, ns		p=.547, ns	

Population Boys –v- Girls: Hyperactivity p<.001; Inattention p<.001; Total Behaviour p<.001.
Sample Boys –v- Girls: Hyperactivity p=.004; Inattention p=.002; Total Behaviour p=.003

Behaviour at baseline

Table 9.2 presents *raw* baseline scores for all measures by gender. Table 9.3 shows behaviour GHA scores over the challenge period by Week and Gender while Table 9.4 presents the same scores by Challenge and Gender.

Table 9.2: Behaviour at baseline by Gender

	TOTAL (n = 144) Median (IQR)	BOYS (n = 75) Median (IQR)	GIRLS (n = 69) Median (IQR)
Teacher	4 (2 to 9)	6 (2 to 14)	3 (1 to 6)
Parent	8 (4 to 14)	11 (7 to 17)	7 (4 to 11.5)
Classroom Obs	35 (24 to 44.75)	40 (30 to 51)	28 (20.5 to 40)
CPT:			
Percent Omissions	4.3 (2.2 to 6.7)	4.5 (2.7 to 7.2)	3.1 (1.6 to 6.5)
Percent Commissions	80.6 (66.7 to 88.9)	88.6 (70.9 to 92.4)	72.2 (61.1 to 83.3)
Raw Score HitRT	386.6 (360.3 to 428.3)	377.1 (346.9 to 414.9)	408.9 (374.9 to 470.8)
Raw Score HitSE	8.8 (6.5 to 12.6)	10.1 (6.6 to 13.9)	8.5 (6.5 to 11.1)
Raw Score Dprime	.16 (-.06 to .31)	.07 (-.09 to .22)	.24 (.03 to .41)
Raw Score Beta	.66 (.49 to .99)	.79 (.60 to 1.14)	.60 (.37 to .90)

Table 9.3: Behaviour GHA by Challenge, Week and Gender

	WEEK 2 Mean (SD)	WEEK 4 Mean (SD)	WEEK 6 Mean (SD)
TOTAL			
	-.06 (.891)	-.002 (.961)	-.63 (.947)
Teacher	-.14 (.985)	-.18 (1.08)	-.34 (.909)
Parent	.13 (1.03)	.30 (1.09)	.09 (1.06)
Classroom	.54 (1.21)	.81 (1.46)	1.08 (1.40)
Obs	.17 (1.01)	.33 (1.09)	.28 (1.00)
CPT			
AGG			
BOYS (n=75)			
Teacher	.19 (1.00)	.32 (1.07)	.24 (1.06)
Parent	.04 (1.00)	.07 (1.19)	-.19 (.945)
Classroom	.42 (1.12)	.696 (1.16)	.46 (1.15)
Obs	.75 (1.28)	1.11 (1.52)	1.3 (1.43)
CPT	.51 (1.07)	.77 (1.15)	.67 (1.03)
AGG			
GIRLS (n=69)			
Teacher	-.34 (.66)	-.36 (.66)	-.395 (.67)
Parent	-.34 (.93)	-.46 (.88)	-.52 (.84)
Classroom	-.198 (.81)	-.14 (.82)	-.31 (.80)
Obs	.31 (1.09)	.47 (1.31)	.795 (1.32)
CPT	-.21 (.77)	-.17 (.77)	-.15 (.76)
AGG			

Table 9.4: Behaviour GHA by Challenge and Gender

	Mix A Mean (SD)	Mix B Mean (SD)	Placebo Mean (SD)
TOTAL SAMPLE	-.07 (.92)	-.02 (.94)	-.03 (.94)
Teacher	-.25 (.94)	-.15 (1.05)	-.26 (.99)
Parent	.17 (.97)	.19 (1.08)	.16 (1.14)
Class Obs	.82 (1.35)	.895 (1.43)	.70 (1.35)
CPT	.25 (1.00)	.33 (1.10)	.19 (1.03)
AGG			
BOYS (n=75)			
Teacher	.21 (1.06)	.30 (1.05)	.24 (1.02)
Parent	-.06 (1.01)	.09 (1.13)	-.09 (1.02)
Class Obs	.46 (1.02)	.61 (1.14)	.51 (1.28)
CPT	1.08 (1.35)	1.18 (1.49)	.90 (1.45)
AGG	.61 (1.00)	.78 (1.15)	.55 (1.11)
GIRLS (n=69)			
Teacher	-.38 (.612)	-.38 (.632)	-.33 (.738)
Parent	-.47 (.818)	-.39 (.917)	-.45 (.930)
Class Obs	-.15 (.804)	-.27 (.784)	-.22 (.824)
CPT	.52 (1.28)	.57 (1.30)	.48 (1.19)
AGG	-.16 (.75)	-.17 (.769)	-.19 (.782)

10. Mixed Model (MM) analyses - Likelihood Ratio Tests

Justification for using an unstructured covariance matrix in the Mixed Models Analyses that follow can be seen by comparing -2LL information criteria of unstructured covariance matrix (Model 1 below) with that of models employing a lesser number of parameters: Model 2 (Compound Symmetry) and Model 3 (AR1). The unstructured model is the best fit model where -2LL figure is smaller indicating a better fit model (Unstructured –v- Compund Symmetry model: p=.0004).

Information Criteria[a]

-2 Restricted Log Likelihood	793.752
Akaike's Information Criterion (AIC)	805.752
Hurvich and Tsai's Criterion (AICC)	805.972
Bozdogan's Criterion (CAIC)	835.534
Schwarz's Bayesian Criterion (BIC)	829.534

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: agg.

1. Fixed Effect = Challenge
Unstructured covariance matrix
(6 parameters)

Information Criteria[a]

-2 Restricted Log Likelihood	813.072
Akaike's Information Criterion (AIC)	817.072
Hurvich and Tsai's Criterion (AICC)	817.103
Bozdogan's Criterion (CAIC)	826.999
Schwarz's Bayesian Criterion (BIC)	824.999

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: agg.

2. Fixed Effect = Challenge
Compound Symmetry
(2 Parameters)

Information Criteria[a]

-2 Restricted Log Likelihood	829.469
Akaike's Information Criterion (AIC)	833.469
Hurvich and Tsai's Criterion (AICC)	833.501
Bozdogan's Criterion (CAIC)	843.397
Schwarz's Bayesian Criterion (BIC)	841.397

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: agg.

3. Fixed Effect = Challenge
Auto Regressive (AR1)
(2 parameters)

11. MM main analyses (1)

A total of 144 children were recruited to enter the study. Of these 14 failed to complete the trial but 136 provided at least one GHA score. These analyses present the results for the Global Hyperactivity Aggregate (GHA) which is an aggregate of the week by week teacher and parent ratings, the CPT composite, and the Classroom Observation Code (COC). The GHA score for any week was calculated when at least 3 out of 4 behaviour measure scores used to calculate the GHA were present. Using Mixed Models to analyse the effects of challenge type (Table 11.1), it was found that for the whole sample there was a significant effect of Mix B in elevating the GHA compared to placebo (p=.013). When the analysis was restricted to those consuming 85% or more of the juice [2] the effects of both Mix A (p=.02) and Mix B (p=.004) were significant (Table 11.1).

11.1: MM ANALYSIS of effect of Challenge on GHA for 8-9yo children with ≥0% [1] and ≥85% juice consumption [2] and complete case analysis [3].

Model	Factor Level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
[1]. (≥0%) Challenge n=136 F(2, 229.04)=3.23, p=.041 -2LL = 793.75	Mix A	132	.25 (.97)	.08 to .42	.25 (.087)	.07 to .42	A –v- P	.08	-.02 to .18	.105
	Mix B	133	.33 (1.10)	.14 to .52	.29 (.087)	.12 to .46	B –v- P	.12	.03 to .22	.013
	Mix P	127	.19 (1.03)	.11 to .37	.17 (.087)	-.01 to .34	-	-	-	-
[2] (≥85%) Challenge n=119 F(2, 175.01)=4.84 p=.009 -2LL = 632.998	Mix A	104	.26 (.93)	.08 to .45	.21 (.091)	.03 to .39	A –v- P	.12	.02 to .23	.02
	Mix B	112	.32 (1.09)	.12 to .52	.24 (.090)	.06 to .42	B –v- P	.15	.05 to .25	.004
	Mix P	103	.19 (1.04)	-.01 to .39	.09 (.092)	-.09 to .27	-	-	-	
[3]: Challenge (complete case) n=91 F(2, 153.90)=5.24 p=.006	Mix A	91	.27 (.92)	.07 to.46	.25 (.104)	.04 to .45	Mix A –v- P	.14	.03 to .24	.015
	Mix B	91	.35 (1.08)	.13 to .58	.28 (.104)	.07 to .49	Mix B –v- P	.17	.06 to .28	.003
	Mix P	91	.19 (1.06)	-.03 to .41	.11 (.104)	-.10 to .32	-	-	-	-

GHA = Global Hyperactivity Aggregate. Est. Marg.Mean = Estimated Marginal Mean

An advantage of the mixed models approach is that it uses a maximum likelihood approach to deal with missing data (Gueorguieva & Krystal , 2004:Mallinckrodt et al, 2004; Molenberghs et al, 2004). This results in optimal and unbiased estimates of the means for each of the effects being tested. As a check that the procedure for dealing with missing values was not producing artefacts in the results a complete case analysis was conducted (see Table 11.1[3] above). This was limited to the children consuming 85% or more of the drinks and with complete GHA values for Mix A, Mix B and Placebo. It was found that the effect sizes for the Mix A vs. Placebo (0.14) and for Mix B vs. Placebo (.17) are similar (in fact slightly larger) to those for children consuming 85% of juice or more.

The challenge periods of central interest were in weeks 2,4 and 6 of the trial when there was a tendency for GHA to increase over time. Further analyses showed that Week of challenge was related to GHA ($F[2,204.02]=7.86$, $p=.001$) but there was no significant Week by Challenge interaction.

To test whether there was any evidence of carry-over effects, the scores of the previous active challenge period were added as factors in the mixed model analysis. Carry-over was also related to GHA ([$F[3,122.60]=5.34$, $p=.002$) but there was no significant Challenge by Carry-over interaction. Pairwise comparisons indicated a significant difference in scores for Mix A, Mix B and Placebo when preceded by 'no treatment' i.e. baseline week, in that all scores preceded by baseline week i.e. Week 2 scores were significantly lower than when order of treatment was preceded either by Placebo ($p=.002$) or Mix A (.042) or Mix B ($p=.009$). However, no significant differences were found between these latter three in terms of order of treatment. Since there was shared variance between Week and Carry-over in that Week 2 scores were equivalent to a 'no treatment' level of Carry-over, only Week was employed in further analyses.

Table 11.2 below shows that the MM analyses adjusted for Week of challenge has had no effect on the analyses presented previously in Table 11.1 although for the group of children who consumed 85% or more of juice in any challenge week the effect of Mix B on the GHA has slightly increased (ES=.16, $p=.001$) and this is also the case for the complete case group (ES=.18, $p=.001$).

Table 11.2: MM Analyses of effect of Challenge on GHA for 8-9yo children with ≥0% [1] and ≥85% juice consumption [2] and complete case analysis [3] adjusted for Week of challenge

Model*	Factor Level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
[1]. (≥0%) Challenge n=136 F(2, 230.84)=3.25, p=.041 -2LL = 793.96	Mix A	132	.25 (.97)	.08 to .42	.29 (.089)	.11 to .46	A –v- P	.08	-.02 to .18	.106
	Mix B	133	.33 (1.10)	.14 to .52	.33 (.089)	.15 to .50	B –v- P	.12	.03 to .22	.013
	Mix P	127	.19 (1.03)	.11 to .37	.21 (.089)	.03 to .38	-	-	-	-
[2] (≥85%) Challenge n=119 F(2, 179.35)=5.57 p=.005 -2LL = 627.30	Mix A	104	.26 (.93)	.08 to .45	.27 (.093)	.08 to .45	A –v- P	.12	.02 to .22	.02
	Mix B	112	.32 (1.09)	.12 to .52	.31 (.093)	.13 to .49	B –v- P	.16	.06 to .26	.001
	Mix P	103	.19 (1.04)	-.01 to .39	.15 (.093)	-.04 to .33	-	-	-	
[3]: Challenge (complete case) n=91 F(2, 157.73)=5.87 p=.003 -2LL = 529.56	Mix A	91	.27 (.92)	.07 to.46	.30 (.106)	.09 to .51	A –v- P	.14	.03 to .24	.014
	Mix B	91	.35 (1.08)	.13 to .58	.34 (.106)	.13 to .55	B –v- P	.18	.07 to .29	.001
	Mix P	91	.19 (1.06)	-.03 to .41	.16 (.106)	-.05 to .37	-	-	-	-

* Adjusted for week of challenge

12. Other possible influencing factors in MM analysis

It was also possible that the effects of the additive mixes would be most marked for those already showing higher GHA scores at baseline. Those with higher Baseline GHA scores had higher GHA scores over the trial period and boys had significantly higher GHA scores than girls (Table 12.1).

Table 12.1: Prior baseline analyses (ANOVA) of possible additional influences

Moderator	N	All Mean (sd)	Groups Mean (sd)	Mean (sd)	Test
			Age ≤106m (n=57)	Age >106m (n=62)	
Age Gp	119	-.02 (1.03)	.07 (1.02)	-.10 (1.03)	F(1,117)=.749, p=.388
			Male (n=64)	Female (n=55)	
Gender	119	-.02 (1.03)	.35 (1.12)	-.44 (.71)	F(1,117)=20.32, p=.000
			Pre-trial diet (n=119)		
Pre-trial diet	119	-.02 (1.03)	-	-	F(1,117)=.035, p=.852
			O level or below	A level or higher	
Maternal education level	117	-.02 (1.03)	.13 (1.17)	-.15 (.90)	F(1,115)=2.02, p=.158

However, when both Baseline GHA and Gender were included in the analysis, the effect of Gender was no longer significant as these measures would have shared variance. Gender was not included in further analyses.

13. MM main analyses (2)

No significance was found for any interactions between Baseline GHA or Week and type of challenge. These interaction terms were therefore not included in further Mixed Model analyses. Table 13.1 below shows analyses of the effect of Challenge on GHA over the study period but adjusted for both Week of challenge and Baseline GHA. The effect of Mix B remains significant in all analyses. However, in the analysis of the group of children with at least 85% consumption of juice in any challenge week, the effect of Mix A on GHA now only approaches significance (p=.063).

13.1: MM Analyses of effect of Challenge on GHA for 8-9yo children with >0% juice consumption [1], >85% juice consumption [2] and a complete case group [3] (adjusting for Baseline GHA and Week of challenge).

Mixed Model*	Factor level	N=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥0%) Challenge F(2, 235.20)=4.00 p=.020 (n=136) -2LL = 615.42	Mix A	132	.25 (.97)	.08 to .42	.27 (.051)	.17 to .37	Mix A –v- P	.08	-.02 to .17	.112
	Mix B	133	.33 (1.10)	.14 to .52	.33 (.051)	.23 to .43	Mix B –v- P	.14	.04 to .23	.005
	Mix P	127	.19 (1.03)	.11 to .37	.19 (.052)	.09 to .29	-	-	-	-
(≥85%) Challenge F(2, 189.32)=4.53 p=.012 (n=119) -2LL = 455.33	Mix A	104	.26 (.93)	.08 to .45	.27 (.052)	-.16 to .37	Mix A –v- P	.09	-.01 to .19	.063
	Mix B	112	.32 (1.09)	.12 to .52	.32 (.051)	.22 to .42	Mix B –v- P	.15	.05 to .25	.003
	Mix P	103	.19 (1.04)	-.01 to .39	.17 (.053)	.07 to .27	-	-	-	-
Challenge F(2, 163.44)=5.61 p=.004 (n=91) -2LL = 392.65	Mix A	91	.27 (.92)	.07 to.46	.29 (.058)	.18 to .41	Mix A –v- P	.12	.02 to .23	.023
	Mix B	91	.35 (1.08)	.13 to .58	.34 (.058)	.23 to .46	Mix B –v- P	.17	.07 to .28	.001
	Mix P	91	.19 (1.06)	-.03 to .41	.17 (.058)	.05 to .28	-	-	-	-

* Adjusted for Baseline GHA and Week of study.

14. MM analyses using all covariates (results report in Lancet (2007) paper)

For the final MM analyses the effect of challenge on GHA was examined again but now with all potential confounds controlled. The results are shown in Table 14.1 for all children, children with at least 85% juice consumption and a complete case analysis. The format of this table is slightly simplified to accommodate parameters for all the covariates.

The results show that for all three analyses there was a statistically significant greater GHA score when challenged with Mix B than with Placebo. For Mix A a significantly greater GHA score compared to Placebo was obtained for the complete case analysis only (ES:.12, .02 to .23, $p<.05$).

Table 14.1. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	.16 [-.01 to .34]	.09 [-.09 to .27]	.11 [-.10 to .32]
Challenge type			
Mix A –v- placebo	.08 [-.02 to .18]	.12 [.02 to .23] *	.14 [.03 to .24] *
Mix B –v- placebo	.12 [.03 to .22] *	.15 [.05 to .25] **	.17 [.06 to .28] **
Model 2			
Intercept	.02 [-.22 to .26]	.14 [-.08 to .37]	.14 [-.12 to .39]
Challenge type			
Mix A –v- placebo	.08 [-.02 to .17]	.09 [-.01 to .19]	.12 [.02 to .23] *
Mix B –v- placebo	.12 [.03 to .22] *	.15 [.05 to .25] **	.17 [.07 to .28] **
Week of study			
Wk 2 –v- Wk 6	-.11 [. 21 to -.00] *	-.19 [-.29 to -.08] **	-.20 [-.32 to -.09] **
Wk 4 –v- Wk 6	.06 [-.03 to .14]	.04 [-.06 to .13]	.03 [-.07 to .13]
Gender	.16 [-.03 to .35]	.08 [-.10 to .26]	.11 [-.09 to .31]
Baseline GHA score	.78 [.69 to .88] ***	.79 [.71 to .88] ***	.79 [.70 to .89] ***
Pre-trial diet	.04 [-.02 to .10]	.03 [-.03 to .09]	.02 [-.05 to .09]
Maternal education level	-.02 [.20 to .16]	-.02 [-.19 to .15]	.01 [-.18 to .21]
Maternal social class	.04 [-.14 to .22]	-.03 [-.20 to .14]	-.06 [-.25 to .13]

* p<.05, ** p<.01, *** p<.001

15. Genotype MM analyses (1)

The following analyses investigate whether the effect of artificial food colors (AFCs) on hyperactivity is moderated by genetic difference between children. Consequently results are presented here for HNMT Thr105Ile, HNMT T939C, COMT Val108Met and ADRA2A c1291G. The effect of each of these polymorphisms on baseline scores was first examined. Table 15.1 indicates no significant effect of these polymorphisms on Baseline GHA.

Table 15.1: Effect of Genotype on Baseline GHA

Moderator	N	All Mean (sd)	Groups Mean (sd)	Mean (sd)	Test
HNMT Thr105Ile	118	-.02 (1.03)	105Ile present (n=24) -.25 (.75)	105Ile absent (n=94) .04 (1.08)	F(1,116)=1.54, p=.217
HNMT T939C	118	-.01 (1.03)	939C present (n=49) .11 (1.19)	939C absent (n=69) -.09 (.89)	F(1,116)=1.03, p=.311
COMT Val108Met	119	-.02 (1.03)	Val108 present (n=85) -.03 (.92)	Val108 absent (n=34) .01 (1.27)	F(1,117)=.046, p=.830
ADRA2A C1291G	116	-.001 (1.03)	1291G present (n=51) -.08 (1.08)	1291G absent (n=65) .06 (1.00)	F(1,114)=.514, p=.475

Analyses were also employed to examine the effect of each of the genetic polymorphisms outlined in Section 4 and their interactions on GHA over the study period. However, these analyses were carried out only for those children with 85% or more consumption of juice. Week of challenge was included in the analyses at this point since this variable was shown to have an impact on GHA over the challenge period.

Tables 15.2 [1 to 4] indicate no significant main effects of these polymorphisms on GHA. However, the tables do provide evidence that the polymorphisms we examined in the HNMT gene moderate the effect of mix A and Mix B on GHA. Tables 14.2 [1] and 14.2 [2] show a significant Challenge (Mix B) by Genotype interaction for both HNMT Thr105Ile (p=.038) and HNMT T939C (p=.036). The Mix A by Genotype interaction also approaches

significance (p=.053) for HNMT T939G. In contrast, there is no evidence that the polymorphisms in the COMT Val108Met or ADRA2A C1291G polymorphisms moderate the effect of either Mix A or Mix B on GHA.

Table 15.2: MM Analyses of effect of Challenge, Week, Genotype (HNMT Thr105Ile), Challenge x HNMT Thr105Ile on GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=118) -2LL = 621.28	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 170.06)=1.11 p=.331	Mix A	103	.27 (.94)	.09 to .45	.17 (.115)	-.06 to .40	Mix A –v- P	.15	.04 to .27	.008
	Mix B	111	.32 (1.09)	.11 to .52	.18 (.114)	-.05 to .40	Mix B –v- P	.21	.10 to .32	.000
	Mix P	102	.19 (1.04)	-.01 to .40	.09 (.114)	-.13 to .32	-	-	-	-
(≥85%) Week F(2, 101.86)=7.42 p=.001	Week 2	105	.14 (.96)	-.05 to .32	.01 (.109)	-.21 to .22	Wk 2 –v- 6	.18	-.29 to .-07	.001
	Week 4	108	.36 (1.10)	.15 to .57	.24 (.118)	.01 to .48	Wk 4 –v- 6	.06	-.03 to .15	.211
	Week 6	103	.28 (1.00)	.08 to .47	.19 (.112)	-.04 to .41	-	-	-	-
(≥85%) HNMT Thr105Ile F(1, 113.80)=2.38 p=.125	105Ile present	24	-.06 (.79)	-.26 to .13	-.02 (.192)	-.40 to .36	105Ile Present –v- Absent	-.20	-.65 to .24	.370
	105Ile absent	94	.35 (1.06)	.22 to .48	.31 (.099)	.12 to .51				
(≥85%) Challenge x HNMT Thr105Ile -105Ile present F(2, 169.92)=2.19 p=.115	Mix A	20	-.06 (.76)	-.41 to .29	-.003 (.204)	-.41 to .40	)	105Ile present –v- 105Ile absent: Mix A –v- P: -.15 (-.40 to .09) p=.221 Mix B –v- P: -.26 (-.50 to -.01) **p=.038**		
	Mix B	24	-.06 (.90)	-.43 to .32	-.05 (.200)	-.45 to .35	)			
	Mix P	22	-.08 (.73)	-.40 to .25	-.004 (.202)	-.40 to .39	)			
(≥85%) Challenge x HNMT Thr105Ile - 105Ile absent	Mix A	83	.35 (.96)	.14 to .56	.34 (.104)	.14 to .55	)			
	Mix B	87	.42 (1.12)	.18 to .66	.40 (.103)	.20 to .61	)			
	Mix P	80	.27 (1.10)	.02 to .51	.19 (.104)	-.02 to .40	)			

Table 15.3: MM Analyses of effect of Challenge, Genotype (HNMT T939C), Challenge x HNMT T939C on GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=118) -2LL = 622.36	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge	Mix A	103	.28 (.92)	.10 to .46	.29 (.094)	.11 to .48	Mix A –v- P	.21	.07 to .34	.002
F(2, 171.82)=4.30 p=.015	Mix B	111	.34 (1.08)	.14 to .54	.33 (.094)	.15 to .52	Mix B –v- P	.26	.13 to .38	.000
	Mix P	103	.19 (1.04)	-.01 to .39	.19 (.094)	.001 to .37	-	-	-	-
(≥85%) Week F(2, 104.29)=8.59 p=.000	Week 2	106	.14 (.96)	-.05 to .32	.12 (.091)	-.06 to .30	Wk 2 –v- 6	-.20	-.30 to -.09	.000
	Week 4	108	.37 (1.09)	.16 to .58	.37 (.101)	.17 to .57	Wk 4 –v- 6	.05	-.04 to .14	.271
	Week 6	103	.30 (.98)	.11 to .49	.32 (.092)	.14 to .50	-	-	-	-
(≥85%) HNMT T939C F(1, 116.07)=1.21 p=.273	939C present	49	.37 (1.11)	.17 to .56	.37 (.135)	.10 to .63	939C present –v- absent	.33	-.03 to .70	.074
	939C absent	69	.20 (.94)	.07 to .34	.17 (.114)	-.05 to .40				

(≥85%) Challenge x HNMT T939C - 939C present F(2, 170.40)=2.73 p=.068	Mix A	41	.37 (1.04)	.04 to .70	.36 (.142)	.07 to .64	)	939C present –v- 939C absent: Mix A –v- P: -.20 (-.41 to .003) **p=.053** Mix B –v- P: -.22 (-.42 to -.01) **p=.036**
	Mix B	47	.43 (1.23)	.07 to .79	.39 (.142)	.11 to .67	)	
	Mix P	43	.29 (1.06)	-.03 to .62	.35 (.142)	.07 to .63	)	
(≥85%) Challenge x HNMT T939C - 939C absent	Mix A	62	.22 (.84)	.01 to .44	.23 (.120)	-.01 to .46	)	
	Mix B	64	.27 (.95)	.03 to .51	.28 (.119)	.04 to .51	)	
	Mix P	60	.11 (1.02)	-.15 to .38	.02 (.121)	-.22 to .26	)	

Table 15.4: MM Analysis of effect of Challenge, Genotype (COMT Val108Met) and Challenge x COMT Val108Met on GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=119) -2LL = 633.74	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge										
F(2, 171.94)=4.27 p=.016	Mix A	104	.26 (.93)	.08 to .45	.25 (.103)	.04 to .45	Mix A –v- P	.12	-.07 to .31	.210
	Mix B	112	.32 (1.09)	.12 to .52	.28 (.103)	.08 to .49	Mix B –v- P	.14	-.05 to .33	.139
	Mix P	103	.19 (1.04)	-.01 to .39	.13 (.104)	-.08 to .33	-	-	-	-
(≥85%) Week	Week 2	106	.14 (.96)	-.05 to .32	.08 (.099)	-.11 to .28	Wk 2 –v- 6	-.18	-.29 to -.07	.001
F(2, 104.36)=7.37 p=.037	Week 4	109	.36 (1.10)	.15 to .57	.31 (.108)	.10 to .53	Wk 4 –v- 6	.05	-.05 to .14	.318
	Week 6	104	.28 (.99)	.09 to .48	.27 (.101)	.07 to .46	-	-	-	-
(≥85%) COMT Val108Met	Val108 present	85	.27 (.99)	.14 to .40	.27 (.104)	.06 to .48	Val108 present –v-absent	.09	-.32 to .49	.665
F(1, 117.22)=.272 p=.603	Val108 absent	34	.24 (1.09)	.01 to .47	.17 (.164)	-.16 to .49				

(≥85%) Challenge x COMT Val108Met - Val108 present F(2, 172.79)=.045 p=.956	Mix A	73	.28 (.89)	.07 to .49	.29 (.110)	.08 to .51	)	Val108 present –v- Val108 absent: Mix A –v- P: .004 (-.22 to .23) p=.975 Mix B –v- P: .03 (-.19 to .25) p=.784
	Mix B	80	.33 (1.03)	.10 to .55	.34 (.109)	.13 to .56	)	
	Mix P	75	.19 (1.05)	-.05 to .44	.17 (.110)	-.05 to .39	)	
(≥85%) Challenge x COMTva- A absent	Mix A	31	.23 (1.03)	-.15 to .61	.20 (.172)	-.14 to .54	)	
	Mix B	32	.31 (1.23)	-.13 to .75	.22 (.172)	-.12 to .56	)	
	Mix P	28	.17 (1.02)	-.22 to .57	.08 (.174)	-.26 to .43	)	

Table 15.5: MM Analysis of effect of Challenge, Week, Genotype (ADRA2A C1291G), Challenge x ADRA2A C1291G on GHA for 8-9yo children with ≥85% juice consumption15. Genotype MM analyses (2)

Mixed Model (n=116) -2LL = 623.89	**Factor level**	**n=**	**GHA unadjusted mean (SD)**	**CI**	**GHA Est. Marg. Mean (SE)**	**CI**	**Comparison main effects**	**Parameter estimate / effect size**	**CI**	**p**
(≥85%) Challenge										
F(2, 174.43)=4.91 p=.008	Mix A	101	.29 (.94)	.10 to .47	.28 (.096)	.09 to .47	Mix A –v- P	.12	-.01 to .26	.077
	Mix B	109	.34 (1.09)	.13 to .55	.32 (.095)	.14 to .51	Mix B –v- P	.17	.03 to .30	**.017**
	Mix P	102	.20 (1.03)	.00 to .41	.17 (.096)	-.02 to .35	-	-	-	-
(≥85%) Week	Week 2	105	.14 (.96)	-.04 to .33	.12 (.092)	-.07 to .30	Wk 2 –v- 6	-.19	-.30 to -.08	**.001**
F(2, 100.09)=7.89 p=.001	Week 4	106	.39 (1.10)	.17 to .60	.35 (.102)	.15 to .55	Wk 4 –v- 6	.05	-.04 to .15	.298
	Week 6	101	.30 (1.00)	.11 to .50	.30 (.094)	.12 to .49	-	-	-	-
(≥85%) ADRA2A C1291G	1291G present	51	.22 (1.02)	.05 to .39	.20 (.134)	-.06 to .47	1291 G present -v- absent	-.09	-.47 to .28	.624
F(1, 112.80)=.343 p=.559	1291G absent	65	.32 (1.03)	.17 to .47	.31 (.119)	.07 to .54				
(≥85%) Challenge x ADRA2A C1291G -G present	Mix A	43	.18 (.90)	-.10 to .46	.22 (.142)	-.06 to .51	)	**1291G present –v- 1291G absent:** Mix A –v- P: -.02 (-.23 to .19) p=.858		
F(2, 174.47)=.018p=.982	Mix B	47	.31 (1.07)	-.01 to .62	.27 (.141)	-.01 to .55	)			
	Mix P	46	.17 (1.08)	-.15 to .49	.12 (.142)	-.16 to .40	)			

(≥85%) Challenge x ADRA2A C1291G - 1291G absent	Mix A Mix B Mix P	58 62 56	.36 (.96) .37 (1.12) .23 (1.00)	.11 to .62 .08 to .65 -.04 to .50	.34 (.126) .38 (.125) .21 (.126)	.09 to .58 .13 to .63 -.04 to .46	)	Mix B –v- P: -.02 (-.22 to .19) p=.883

16. Genotype MM analysis (2)

With the inclusion of Baseline GHA in the MM main analyses (2), the addition of Gender to the model did not have any added effect since Baseline GHA and the effect of Gender on behaviour would have shared variance. However, gender has been shown to moderate the effects of genotype in a number of ways over a number of conditions in children and for this reason the above Genotype analyses were re-examined with the inclusion of Gender and a Genotype by Gender interaction term. This was carried out in order to check if the interaction term accounted for variance over and above that accounted for by Baseline GHA. In no case was the interaction term found to be significant and these terms were therefore omitted from the genotype MM analyses that follow.

17. Genotype MM analyses (3)

Previous Genotype analyses indicate that both the HNMT Thr105Ile and HNMT T939C polymorphisms play a possible moderating role in the relationship between Challenge and GHA. Certainly in the case of Mix B, those children with the 105Ile absent genotype in the HNMT Thr105Ile polymorphism and those with 939C absent genotype for the HNMT T939C polymorphism may be more vulnerable to the effects of additives in their food . Further analyses are presented below for each of these polymorphisms, but including both Week of challenge and Baseline GHA and based only on children with ≥85% juice consumption.

Tables 17.1 and 17.2 show that the inclusion of both Week and Baseline GHA have slightly changed findings. The main effects of both Mix A and Mix B on GHA in analyses remain significant. However, the difference between HNMT 105Ile present and absent genotypes now approaches significance for Mix B compared to Mix P (p=.053) with those children with T absent alleles showing higher scores (Table 17.1). The Mix A –v- P comparison remains non-significant (p=.402). However, in the case of HNMT T939C, the difference between 939C present and absent genotypes is now significant for both Mix A compared to Mix P (p=.023) and Mix B compared to Mix P (p=.03) with for both challenges those with the 939C absent genotype showing significantly higher scores (Table 17.2). Overall, these final set of Genotype analyses again indicate that both the HNMT Thr105Ile and HNMT T939C polymorphisms may play a possible moderating role in the relationship between Challenge and GHA. This is certainly so in the case of the HNMT T939C where those children with the C absent allele appear to be particularly vulnerable to the effects of additives in their food compared to those with the allele present.

Table 17.1: MM Analysis of effect of Challenge, Week, Genotype (HNMT Thr105Ile), Challenge x HNMT Thr105Ile and Baseline GHA on GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=118) -2LL = 453.90	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 179.42)=.815 p=.444	Mix A	103	.27 (.94)	.09 to .45	.24 (.065)	.11 to .37	Mix A –v- P	.11	.003 to .23	.044
	Mix B	111	.32 (1.09)	.11 to .52	.25 (.062)	.13 to .37	Mix B –v- P	.19	.08 to .30	.001
	Mix P	102	.19 (1.04)	-.01 to .40	.18 (.063)	.05 to .30	-	-	-	-
(≥85%) Week F(2, 102.16)=7.18 p=.001	Week 2	105	.14 (.96)	-.05 to .32	.09 (.053)	-.02 to .19	Wk 2 –v- 6	-.18	-.29 to .-07	.001
	Week 4	108	.36 (1.10)	.15 to .57	.31 (.069)	.18 to .45	Wk 4 –v- 6	.05	-.04 to .14	.295
	Week 6	103	.28 (1.00)	.08 to .47	.27 (.061)	.14 to .39	-	-	-	-
(≥85%) Baseline GHA F(1, 108.11)=408.17 p=.000		-	-	-	-	-	-	.80	.72 to .87	.000
(≥85%) HNMT Thr105Ile	105Ile present	24	-.06 (.79)	-.26 to .13	.17 (.091)	-.01 to .35	105Ile present	.004	-.24 to .24	.972

F(1, 105.99)=1.19 p=.151	105Ile absent	94	.35 (1.06)	.22 to .48	.28 (.048)	.18 to .37	–v- absent	
(≥85%) Challenge x HNMT Thr105Ile -105Ile present F(2, 179.21)=1.91 p=.151	Mix A	20	-.06 (.76)	-.41 to .29	.19 (.114)	-.03 to .42	)	
	Mix B	24	-.06 (.90)	-.43 to .32	.13 (.108)	-.08 to .34	))	105Ile present –v- 105Ile absent:
	Mix P	22	-.08 (.73)	-.40 to .25	.18 (.109)	-.04 to .39	))	Mix A –v- P: -.10 (-.35 to .14) p=.402
(≥85%) Challenge x HNMT Thr105Ile - 105Ile absent	Mix A	83	.35 (.96)	.14 to .56	.29 (.058)	.18 to .41		Mix B –v- P: -.24 (-.48 to .003) **p=.053**
	Mix B	87	.42 (1.12)	.18 to .66	.36 (.057)	.25 to .48	)	
	Mix P	80	.27 (1.10)	.02 to .51	.18 (.059)	.06 to .29	))	

Table 17.2: MM Analysis of effect of Challenge,Genotype (HNMT T939C), Challenge x HNMT T939C and Baseline GHA on GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=118) -2LL = 454.04	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 182.31)=3.58 p=.030	Mix A	103	.28 (.92)	.10 to .46	.27 (.053)	.16 to .37	Mix A –v- P	.19	.06 to .32	.003
	Mix B	111	.34 (1.08)	.14 to .54	.33 (.052)	.22 to .43	Mix B –v- P	.24	.12 to .37	.000
	Mix P	103	.19 (1.04)	-.01 to .39	.19 (.053)	.09 to .30	-	-	-	-
(≥85%) Week F(2, 104.49)=8.74 p=.000	Week 2	106	.14 (.96)	-.05 to .32	.11 (.045)	.02 to .20	Wk 2 –v- 6	-.20	-.31 to -.10	.000
	Week 4	108	.37 (1.09)	.16 to .58	.36 (.063)	.23 to .48	Wk 4 –v- 6	.04	-.05 to .13	.382
	Week 6	103	.30 (.98)	.11 to .49	.32 (.053)	.21 to .42	-	-	-	-
(≥85%) Baseline GHA F(1, 108.61)=418.29 p=.000	-	-	-	-	-	-	-	.80	.72 to .88	.000
(≥85%) HNMT T939C F(1, 110.26)=.01 p=.920	939C present	49	.37 (1.11)	.17 to .56	.27 (.065)	.14 to .40	939C present –v- absent	.16	.72 to .88	.114
	939C absent	69	.20 (.94)	.07 to .34	.26 (.055)	.15 to .37				

(≥85%) Challenge x HNMT T939C - 939C present F(2, 180.96)=3.32 p=.038	Mix A	41	.37 (1.04)	.04 to .70	.23 (.080)	.07 to .39	)	939C present –v- 939C absent: Mix A –v- P: -.23 (-.44 to -.03) **p=.023** Mix B –v- P: -.22 (-.42 to -.02) **p=.030**	
	Mix B	47	.43 (1.23)	.07 to .79	.30 (.078)	.14 to .45	))		
	Mix P	43	.29 (1.06)	-.03 to .62	.27 (.078)	.12 to .43			
(≥85%) Challenge x HNMT T939C - 939C absent	Mix A	62	.22 (.84)	.01 to .44	.31 (.067)	.17 to .44	)		
	Mix B	64	.27 (.95)	.03 to .51	.36 (.065)	.23 to .49	)		
	Mix P	60	.11 (1.02)	-.15 to .38	.11 (.068)	-.02 to .25	)		

18. Genotype MM analyses using all covariates

The analyses of the moderating effects of genotype were run again but on this occasion using all the potential confound that had been included in the analyses in Section 14. The results are presented in the same reduced form as was adopted in Section 14 to provided parameters on all the covariates. The results are shown in Table 18.1 to Table 18.2 for children with at least 85% juice consumption for the HNMT gene polymorphisms.

It can be seen in Table 18.1 that with the full set of covariates that there is no general effect of HNMT Thr105Ile on the GHA.. The adverse effect of Mix A compared to Placebo is not significantly moderated by the HNMT Thr105Ile polymorphism . The effect of Mix B compared to Placebo is significantly greater for the HNMT 105Ile absent genotype than for the 105Ile present genotype (ES: -.24, -.48 to -.00 p=.050). In Table 18.2 there is no effect of the HNMT T939C polymorphisms on the overall GHA. The pattern of results for HNMT T939C is that for 939C absent the effects of Mix A compared to Placebo are greater than for 939C present (ES: -.24, -.44 to -.04,p=.021) and similarly for Mix B compared to Placebo (ES:-.23, -.43 to -.03, p=.026).

Table 18.1. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (HNMT Thr105Ile) and Challenge type by Genotype interaction for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	**≥85% consumption**
Model 1	
Intercept	.12 (-.08 to .33), p=.232
Challenge type	
Mix A –v- placebo	.15 (.04 to .27), p=.009
Mix B –v- placebo	.20 (.09 to .31), p=.001
Genotype – HNMT Thr105Ile 105Ile present –v- 105Ile absent	-.20 (-.65 to .24), p=.370
Challenge x HNMT Thr105Ile 105Ile present –v- 105Ile absent	
Mix A –v- placebo	-.14 (-.39 to .11), p=.270
Mix B –v- placebo	-.25 (-.50 to -.002), p=.048
Model 2	
Intercept	.14 (-.10 to .38), p=.253
Challenge type	
Mix A –v- placebo	.11 (.002 to .23), p=.046
Mix B –v- placebo	.19 (.08 to .30), p=.001
HNMT Thr105Ile 105Ile present –v- 105Ile absent	.01 (-.24 to .26), p=.956
Challenge x HNMT Thr105Ile 105Ile present –v- 105Ile absent	
Mix A –v- placebo	-.10 (-.35 to .14), p=.403
Mix B –v- placebo	-.24 (-.48 to .00), p=.050
Week of study	
Wk 2 –v- Wk 6	-.18 (-.29 to -.07), p=.001
Wk 4 –v- Wk 6	.05 (-.04 to .14), p=.299
Gender	.09 (-.09 to .27), p=.331
Baseline GHA score	.78 (.70 to .87), p<.001
Pre-trial diet	.02 (-.04 to .08), p=.536
Maternal education level	-.00 (-.18 to .17), p=.983
Maternal social class	-.03 (-.20 to .14), p=.738

Table 18.2. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (HNMT T939C) and Challenge type by Genotype interaction for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	**≥85% consumption**
Model 1	
Intercept	-.03 (-.27 to .21), p=.814
Challenge type	
Mix A –v- placebo	.19 (.06 to .33), p=.004
Mix B –v- placebo	.23 (.10 to .36), p=.001
Genotype – HNMT T939C 939C present –v- 939C absent	.31 (-.06 to .67), p=.097
Challenge x HNMT T939C 939C present –v- 939C absent	
Mix A –v- placebo	-.17 (-.38 to .04), p=.106
Mix B –v- placebo	-.18 (-.39 to .02), p=.084
Model 2	
Intercept	.08 (-.16 to .33), p=.512
Challenge type	
Mix A –v- placebo	.19 (.07 to .32), p=.003
Mix B –v- placebo	.25 (.12 to .37), p<.001
Genotype – HNMT T939C 939C present –v- 939C absent	.18 (-.03 to .38), p=.089
Challenge x HNMT T939C 939C present –v- 939C absent	
Mix A –v- placebo	-.24 (-.44 to -.04), p=.021
Mix B –v- placebo	-.23 (-.43 to -.03), p=.026
Week of study	
Wk 2 –v- Wk 6	-.21 (-.31 to -.10), p<.001
Wk 4 –v- Wk 6	.04 (-.05 to .13), p=.380
Gender	.08 (-.09 to .26), p=.357
Baseline GHA score	.79 (.70 to .87), p<.001
Pre-trial diet	.02 (-.04 to .08), p=.416
Maternal education level	-.03 (-.20 to .14), p=.729
Maternal social class	-.03 (-.20 to .14), p=.747

Table 18.3. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (COMTva) and Challenge type by Genotype interaction for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	**≥85% consumption**
Model 1	
Intercept	.05 (-.29 to .39), p=.762
Challenge type	
Mix A –v- placebo	.11 (-.08 to .30), p=.257
Mix B –v- placebo	.10 (-.09 to .29), p=.323
Genotype – COMT Val108Met Val108 present –v- Val108 absent	.05 (-.36 to .45), p=.819
Challenge x COMT Val108Met Val108 present –v- Val108 absent	
Mix A –v- placebo	.02 (-.20 to .25), p=.842
Mix B –v- placebo	.08 (-.15 to .30), p=.491
Model 2	
Intercept	.03 (-.26 to .33), p=.821
Challenge type	
Mix A –v- placebo	.08 (-.10 to .27), p=.379
Mix B –v- placebo	.14 (-.05 to .32), p=.151
Genotype – COMT Val108Met Val108 present –v- Val108 absent	.12 (-.11 to .35), p=.295
Challenge x COMT Val108Met Val108 present –v- Val108 absent	
Mix A –v- placebo	.02 (-.20 to .24), p=.874
Mix B –v- placebo	.02 (-.20 to .24), p=.865
Week of study	
Wk 2 –v- Wk 6	-.19 (-.30 to -.08), p=.001
Wk 4 –v- Wk 6	.04 (-.06 to .13), p=.447
Gender	.11 (-.07 to .29), p=.246
Baseline GHA score	.79 (.70 to .88), p<.001
Pre-trial diet	.03 (-.03 to .09), p=.335
Maternal education level	-.02 (-.19 to .15), p=.823
Maternal social class	-.03 (-.20 to .14), p=.698

Table 18.4. General Linear Mixed Models estimates for GHA over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo, Genotype (ADRDA2a) and Challenge type by Genotype interaction for 8/9 year old children (Model 1) and with the effects of week during study, gender, GHA in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)
Explanatory variable	**≥85% consumption**
Model 1	
Intercept	.16 (-.09 to .40), p=.214
Challenge type	
Mix A –v- placebo	.12 (-.02 to .26), p=.083
Mix B –v- placebo	.15 (.01 to .29), p=.032
Genotype – ADRDA2A C1291G 1291G present –v- 1291G absent	-.10 (-.47 to .27), p=.585
Challenge x ADRDA2A C1291G 1291G present –v- 1291G absent	
Mix A –v- placebo	-.01 (-.22 to .21), p=.944
Mix B –v- placebo	-.01 (-.22 to .20), p=.918
Model 2	
Intercept	-12 (-.15 to .39), p=.393
Challenge type	
Mix A –v- placebo	.11 (-.02 to .25), p=.104
Mix B –v- placebo	.14 (.01 to .28), p=.036
Genotype - ADRDA2A C1291G 1291G present –v- 1291G absent	.05 (-.16 to .26), p=.649
Challenge x ADRDA2A C1291G 1291G present –v- 1291G absent	
Mix A –v- placebo	-.05 (-.26 to .15), p=.607
Mix B –v- placebo	-.004 (-.21 to .20), p=.967
Week of study	
Wk 2 –v- Wk 6	-.19 (-.30 to -.08), p=.001
Wk 4 –v- Wk 6	.04 (-.06 to .13), p=.416
Gender	.07 (-.11 to .26), p=.429
Baseline GHA score	.79 (.71 to .88), p<.001
Pre-trial diet	.03 (-.03 to .09), p=.316
Maternal education level	-.02 (-.20 to .16), p=.850
Maternal social class	-.03 (-.20 to .15), p=.763

19. Dietary infractions (no ‘Whoops’) analysis

A further analysis of interest was the effect of Challenge on GHA only for those children with no ‘Whoops’ recorded over the period of the study (n=107 children). This was completed only for those children with 85% or more juice consumption in any challenge week and no recorded Whoops. Both Week of challenge and Baseline GHA were included in analysis. Table 19.1 shows a significant effect of Mix B on GHA. Juice was kept in a fridge and consumed at home either prior to the child’s attendance at school but mainly on return from school. Parents were therefore better able to keep a record of juice consumed. However, they relied on their own children to some extent to report on food consumed through the school day and this may not have been as accurately recorded in some cases as juice consumption at home.

Table 19.1: MM Analysis of effect of Challenge, Week and Baseline GHA on GHA for 8-9yo children with ≥85% juice consumption and no recorded 'Whoops'

Mixed Model (n=107) -2LL = 333.68	Factor level	N=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 133.02)=4.16 p=.018	Mix A	76	.29 (.96)	.07 to .51	.23 (.059)	.12 to .35	Mix A –v- P	.04	.-08 to .16	.505
	Mix B	75	.31 (1.12)	.05 to .57	.36 (.059)	.24 to .47	Mix B –v- P	.16	.05 to .28	.006
	Mix P	81	.21 (1.02)	-.02 to .43	.19 (.057)	.08 to .30	-	-	-	-
(≥85%) Week F(2, 73.68)=7.52 p=.001	Week 2	79	.12 (.97)	-.10 to .34	.12 (.053)	.02 to .23	Wk 2 –v- 6	-.16	-.29 to -.04	.009
	Week 4	79	.36 (1.09)	.11 to .60	.37 (.064)	.24 to .50	Wk 4 –v- 6	.09	-.03 to .20	.136
	Week 6	74	.33 (1.01)	.09 to .56	.29 (.059)	.17 to .40	-	-	-	-
(≥85%) Baseline GHA F(1, 96.96)=331.33 p=.000	-	-	-	-	-	-	-	.80	.72 to .89	.000

20. Disaggregated measures analyses

The GHA score represented an unweighted average of the standardised Parent, Teacher, Classroom Observation and Continuous Performance Task behaviour scores obtained over weeks of challenge. Further additional analyses of interest included the effect of Challenge on the disaggregated standardised GHA behaviour scores for each of the behaviour measures.

Both Week and Baseline GHA were again included in these analyses. In each case, however, the Baseline GHA was that for the particular measure under analysis so that, for example, the parent Baseline GHA was included in the MM analysis of the effect of Challenge on Parent GHA.

These analyses were carried out to help highlight the individual components of GHA over the challenge period which may have contributed to any relationship between consumption of additives and children's behaviour. However, it should be noted that each behaviour measure employed in the study focuses on slightly differing aspects of hyperactive behaviour over a range of differing contexts. The rationale, in the first instance, for using an aggregated measure (GHA) to record behaviour in this study was to obtain a behaviour score which reflected the child's behaviour over all of these differing situations. In addition, additive mixes were consumed at home and not at school and any behaviour score within a particular context will be a function, amongst other things, not only of the amount of juice consumed but also time and individual differences in absorption of additives. Findings presented in Tables 20.1 to 20.4 below should be viewed in this light.

Analyses show a significant effect of Mix B ($p=.044$) on GHA based on parental scores but no significant effect of Mix A. It may be that parents were more sensitive to changes in behaviour apparent at home following consumption of juice.

Table 20.1: MM Analysis of effect of Challenge, Week and Baseline Parent GHA on Parent GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=119) -2LL=580.75	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
Challenge F(2, 196.78)=2.27 p=.106	Mix A	104	-.22 (.94)	-.40 to -.03	-.22 (.062)	-.34 to -.10	Mix A –v- P	.03	-.10 to .16	.642
	Mix B	113	-.09 (1.09)	-.30 to .11	-.12 (.060)	-.24 to -.01	Mix B –v- P	.13	.003 to .25	.044
	Mix P	103	-.30 (1.00)	-.49 to -.10	-.25 (.062)	-.37 to -.13	-	-	-	-
Week F(2, 99.73)=4.72 p=.011	Week 2	107	.-12 (1.00)	-.32 to .07	-.13 (.056)	-.24 to -.02	Wk 2 –v- 6	.18	.04 to .33	.012
	Week 4	109	-.13 (1.11)	-.34 to .08	-.16 (.066)	-.29 to -.03	Wk 4 –v- 6	.15	.04 to .27	.009
	Week 6	104	-.35 (.90)	-.53 to -.17	-.31 (.064)	-.44 to -.18	-	-	-	-
Baseline Parent GHA F(1, 124.96)=269.15 p=.000	-	-	-	-	-	-	-	.75	.66 to .85	.000

Table 20.2: MM Analysis of effect of Challenge, Week and Baseline Teacher GHA on Teacher GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=119) -2LL = 437.44	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
Challenge F(2, 203.96)=.049 p=.952	Mix A	104	-.03 (.95)	-.21 to .16	-.04 (.050)	-.14 to .05	Mix A –v- P	-.01	-.11 to .09	.834
	Mix B	112	-.03 (.97)	-.21 to .15	-.03 (.049)	-.13 to .07	Mix B –v- P	.005	-.09 to .10	.928
	Mix P	103	-.05 (.93)	-.23 to .13	-.03 (.050)	-.13 to .07				
Week F(2, 105.93)=1.62 p=.203	Week 2	106	-.06 (.90)	-.23 to .12	-.08 (.040)	-.16 to .005	Wk 2 –v- 6	-.02	-.13 to .08	.651
	Week 4	109	.03 (1.00)	-.16 to .22	.02 (.059)	-.10 to .14	Wk 4 –v- 6	.07	-.03 to .17	.183
	Week 6	104	-.08 (.95)	-.27 to .10	-.05 (.051)	-.15 to .05	-	-	-	-
Baseline Teacher GHA F(1, 115.83)=470.22 p=.000	-	-	-	-	-	-	-	.77	.70 to .84	.000

Table 20.3: MM Analysis of effect of Challenge, Week and Baseline COC GHA on Classroom Observation GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=119) -2LL = 664.37	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 197.74)=.606 p=.547	Mix A	104	.21 (1.01)	.02 to .41	.20 (.072)	.06 to .35	Mix A –v- P	.08	-.06 to .22	.274
	Mix B	112	.18 (1.06)	-.02 to .38	.17 (.070)	.03 to .31	Mix B –v- P	.05	-.10 to .19	.518
	Mix P	103	.13 (1.17)	-.09 to .36	.12 (.072)	-.02 to .27				
(≥85%) Week F(2, 106.43)=7.03 p=.001	Week 2	106	.09 (1.01)	-.11 to .28	.07 (.068)	-.07 to .20	Wk 2 –v- 6	-.04	-.20 to .11	.559
	Week 4	109	.34 (1.11)	.13 to .55	.32 (.078)	.17 to .48	Wk 4 –v- 6	.21	.08 to .35	.002
	Week 6	104	.09 (1.09)	-.12 to .30	.11 (.070)	-.03 to .25	-	-	-	-
Baseline COC GHA F(1, 108.40)=174.88 p=.000	-	-	-	-	-	-	-	.71	.60 to .81	.000

20.4: MM Analysis of effect of Challenge, Week and Baseline CPT GHA on CPT GHA for 8-9yo children with ≥85% juice consumption

Mixed Model (n=94) -2LL = 730.55	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	p
(≥85%) Challenge F(2, 155.12)=1.41 p=.249	Mix A	82	.-.10 (.93)	-.31 to .10	.79 (.127)	.54 to 1.04	Mix A –v- P	.08	-.16 to .32	.509
	Mix B	89	-.09 (.92)	-.28 to .11	.91 (.124)	.67 to 1.16	Mix B –v- P	.20	-.04 to .43	.099
	Mix P	82	-.05 (.85)	-.24 to .13	.71 (.127)	.46 to .96				
(≥85%) Week F(2, 84.23)=10.63 p=.000	Week 2	83	-.07 (.86)	-.26 to .11	.51 (.109)	.30 to .73	Wk 2 –v- 6	-.54	-.77 to -.31	.000
	Week 4	87	-.11 (.92)	-.30 to .09	.85 (.137)	.58 to 1.13	Wk 4 –v- 6	-.19	-.44 to .05	.119
	Week 6	83	-.06 (.92)	-.26 to .14	1.05 (.134)	.78 to 1.31	-	-	-	-
Baseline CPT GHA F(1, 92.74)=37.17 p=.000	-	-	-	-	-	-	-	.69	.46 to .91	.000

21. Disaggregated measures analyses using all covariates

In order to provide parallel information to that for the effect of challenge on GHA the disaggregated measures were examined again but now with all potential confounds controlled. The results are shown in Table 21.1 for all children, children with at least 85% juice consumption and a complete case analysis. The format of this table is slightly simplified to accommodate parameters for all the covariates.

Any single indicator is likely to be relatively less reliable compared to the aggregate measure. The consequent increased measurement error makes it less likely that a significant effect will be detected. For this reason the results are most appropriately discussed in terms of the effect sizes. For Challenge type under Model 2 20 of the 24 effect sizes in Tables 21.1 to 21.4 are in the direction of hyperactivity being more marked under the active than the placebo challenge. The strongest effects are found for the computerised test of attention (CPTII).

Table 21.1. General Linear Mixed Models estimates for **parent reported behaviou**r over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **8/9 year old** children (Model 1) and with the effects of week during study, gender, parent reported behaviour in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	-.31 [-.48 to -.15] p <.001	-.30 [-.48 to -.12] p =.001	-.30 [-.50 to -.10] p =.004
Challenge type			
Mix A –v- placebo	.03 [-.09 to .14] p =.666	.05 [-.08 to .17] p =.492	.04 [-.08 to .17] p =.493
Mix B –v- placebo	.10 [-.01 to .22] p =.081	.11 [-.02 to .24] p =.097	.06 [-.07 to .19] p =.365
Model 2			
Intercept	-.31 [-.57 to -.05] p=.020	-.26 [-.53 to .004] p=.053	-.24 [-.54 to .06] p=.120
Challenge type			
Mix A –v- placebo	.01 [-.10 to .13] p=.821	.03 [-.10 to .16] p=.663	.03 [-.09 to .16] p=.581
Mix B –v- placebo	.13 [.01 to .24] p=.031	.13 [.002 to .25] p=.046	.08 [-.05 to .20] p=.237
Week of study			
Wk 2 –v- Wk 6	.18 [.05 to .31] p=.008	.18 [.04 to .32] p=.014	.20 [.06 to .35] p=.008
Wk 4 –v- Wk 6	.16 [.06 to .26] p=.003	.16 [.04 to .27] p=.009	.13 [.02 to .24] p=.024
Gender	-.05 [-.24 to .15] p=.624	-.12 [-.31 to .08] p=.228	-.06 [-.28 to .17] p=.600
Baseline GHA score	.73 [.64 to .83] p<.001	.79 [.69 to .88] p<.001	.79 [.68 to .90] p<.001
Pre-trial diet	-.00 [-.07 to .06] p=.930	.02 [-.05 to .09] p=.611	.002 [-.08 to .08] p=.966
Maternal education level	-.15 [-.35 to .05] p=.129	-.19 [-.39 to .01] p=.064	-.18 [-.40 to .04] p=.116
Maternal social class	.06 [-.13 to .25] p=.551	.03 [-.17 to .22] p=.793	.01 [-.22 to .23] p=.960

Table 21.2. General Linear Mixed Models estimates for **teacher reported behaviour** over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **8/9 year old** children (Model 1) and with the effects of week during study, gender, teacher reported behaviour in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	-.02 [-.18 to .13] p =.766	-.07 [-.25 to .10] p =.393	-.08 [-.28 to .12] p =.419
Challenge type			
Mix A –v- placebo	-.04 [-.13 to .06] p =.432	.01 [.-10 to .11] p =.898	.02 [-.10 top .13] p =.762
Mix B –v- placebo	-.01 [-.11 to .08] p =.765	.03 [-.08 to .13] p =.631	.04 [-.07 to .16] p =.439
Model 2			
Intercept	-.23 [-.43 to -.02] p=.028	-.20 [-.41 to .004] p=.054	-.17 [-.40 to .06] p=.145
Challenge type			
Mix A –v- placebo	-.04 [-.13 to .06] p=.453	-.01 [-.12 to .09] p=.800	-.004 [-.11 to .10] p=.941
Mix B –v- placebo	-.03 [-.13 to .06] p=.514	.01 [-.09 to .11] p=.848	.04 [-.07 to .14] p=.507
Week of study			
Wk 2 –v- Wk 6	.003 [-.09 to .10] p=.944	-.02 [-.13 to .08] p=.667	-.03 [-.15 to .08] p=.593
Wk 4 –v- Wk 6	.05 [-.04 to .15] p=.296	.07 [-.03 to .17] p=.181	.06 [-.05 to .17] p=.278
Gender	.12 [-.03 to .28] p=.127	.12 [-.03 to .28] p=.118	.18 [.00 to .35] p=.045
Baseline GHA score	.73 [.66 to .81] p<.001	.75 [.68 to .82] p<.001	.78 [.69 to .86] p<.001
Pre-trial diet	.02 [-.04 to .07] p=.520	.01 [-.04 to .07] p=.651	-.01 [-.07 to .05] p=.695
Maternal education level	.11 [-.05 to .26] p=.179	.05 [-.10 to .21] p=.496	.05 [-.12 to .22] p=.541
Maternal social class	.08 [-.07 to .24] p=.281	.03 [-.12 to .18] p=.699	.02 [-.15 to .18] p=.856

Table 21.3. General Linear Mixed Models estimates for **classroom observation** over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **8/9 year old** children (Model 1) and with the effects of week during study, gender, teacher reported behaviour in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	.12 [-.06 to .30] p =.195	.07 [-.13 to .27] p =.483	.09 [-.13 to .31] p =.425
Challenge type			
Mix A –v- placebo	.04 [-.09 to .17] p =.582	.10 [-.05 to .25] p =.208	.07 [-.09 to .23] p =.397
Mix B –v- placebo	.02 [-.11 to .15] p =.771	.07 [-.08 to .22] p =.353	.08 [-.08 to .24] p =.322
Model 2			
Intercept	-.18 [-.48 to .12] p=.232	-.08 [-.39 to .22] p=.586	-.09 [-.42 to .25] p=.610
Challenge type			
Mix A –v- placebo	.02 [-.11 to .15] p=.726	.08 [-.07 to .22] p=.298	.04 [-.11 to .20] p=.593
Mix B –v- placebo	.01 [-.12 to .14] p=.886	.05 [-.09 to .19] p=.473	.07 [-.09 to .22] p=.390
Week of study			
Wk 2 –v- Wk 6	.04 [-.10 to .18] p=.594	-.05 [-.20 to .10] p=.516	-.06 [-.23 to .10] p=.455
Wk 4 –v- Wk 6	.23 [.11 to .35] p<.001	.21 [.08 to .35] p=.002	.20 [.06 to .35] p=.007
Gender	.35 [.12 to .57] p=.003	.34 [.11 to .57] p=.004	.40 [.15 to .65] p=.002
Baseline GHA score	.68 [.57 to .79] p<.001	.67 [.56 to .78] p<.001	.64 [.53 to .74] p<.001
Pre-trial diet	.02 [-.06 to .10] p=.623	-.01 [-.10 to .07] p=.730	-.04 [-.13 to .05] p=.342
Maternal education level	.02 [-.21 to .25] p=.861	-.02 [-.26 to .21] p=.842	.06 [-.19 to .31] p=.628
Maternal social class	-.03 [-.26 to .19] p=.787	-.03 [-.26 to .20] p=.793	-.02 [-.27 to .22] p=.844

Table 21.4. General Linear Mixed Models estimates for **CPTII aggregate** over challenge period for Additive Mix A versus Placebo and Additive Mix B versus Placebo for **8/9 year old** children (Model 1) and with the effects of week during study, gender, CPTII aggregate in baseline week, number of additives in pre-trial diet, maternal educational level and social class controlled (Model 2).

	Estimate (95% CI)		
Explanatory variable	**Whole sample** n=136	**≥85% consumption** n=119	**Complete case** n=91
Model 1			
Intercept	.61 [.38 to .84] p <.001	.50 [.26 to .73] p <.001	.52 [.27 to .78] p <.001
Challenge type			
Mix A –v- placebo	.13 [-.08 to .35] p =.223	.12 [-.11 to .35] p =.298	.20 [-.05 to .44] p =.112
Mix B –v- placebo	.18 [-.03 to .39] p =.100	.14 [-.09 to .36] p =.225	.23 [-.01 to .48] p =.061
Model 2			
Intercept	.68 [.14 to 1.22] p=.015	.82 [.28 to 1.36] p=.003	.76 [.14 to 1.38] p=.017
Challenge type			
Mix A –v- placebo	.10 [-.13 to .34], p=.393	.08 [-.16 to .32] p=.509	.18 [-.08 to .44] p=.163
Mix B –v- placebo	.19 [-.05 to .42] p=.118	.20 [-.04 to .43] p=.100	.32 [.06 to .59] p=.015
Week of study			
Wk 2 –v- Wk 6	-.44 [-.67 to -.20] p<.001	-.54 [-.77 to -.31] p<.001	-.55 [-.80 to -.30] p<.001
Wk 4 –v- Wk 6	-.13 [-.36 to .11] p=.282	-.20 [-.45 to .05] p=.110	-.17 [-.44 to .10] p=.221
Gender	.23 [-.20 to .66] p.292	.13 [-.30 to .55] p=.551	.06 [-.43 to .55] p=.811
Baseline GHA score	.66 [.41 to .90] p<.001	.69 [.44 to .93] p<.001	.60 [.31 to .89] p<.001
Pre-trial diet	.08 [-.06 to .23] p=.270	.06 [-.09 to .21] p=.429	.08 [-.09 to .26] p=.349
Maternal education level	.03 [-.40 to .46] p=.899	.13 [-.30 to .55] p=.560	.12 [-.36 to .61] p=.613
Maternal social class	-.06 [-.49 to .37] p=.786	-.16 [-.59 to .27] p=.470	-.13 [-.63 to .36] p=.598

22. Responders and Non-responders to additives analyses

An additional question of interest relates to responders and non-responders to the additives and whether, for example, the observed differences in responses to Mix A and Mix B in this report might be the result of a similar group of children tending to respond to both mixes or due to some individuals tending to respond more to one challenge compared to the other.

For those children who completed the study, the difference scores between Mix A and Mix P (mean .04, SD .68 [IQR -.35 to .38]) and those for Mix B and Mix P (mean .13, SD .62 [IQR -.23 to .50]) .were calculated. Those children with scores on or above the 75th percentile were classed as 'responders', those on or below the 25th percentile as 'non-responders', and the remaining children as 'neutral' responders. Of the 31 children who were responders to Mix A and 32 children who were responders to Mix B, 15/48 (31.3%) of responders to Mix A and/or Mix B were responders to both mixes, 16 (33.3%) were responders to Mix A only and 17 were responders to Mix B only (35.4%). It is not clear as yet what characteristics of the responders resulted in a stronger response to Mix B compared to Mix A in the main analyses.

23. Acute Challenge

The aim of this Stage Two Acute Challenge phase of the study was to allow by separate challenge examination not only of neuropsychological and behavioural responses to additives in food but also to explore their relationship to metabolic factors that may mediate such responses across time. It is thought that additives in food may act to produce a pharmacological effect mediated by histamine release. In addition to behavioural measures, in this stage of the study, urine samples were collected to test for histamine using urinary creatinine as a standard for measurement. Saliva samples were also collected to be assayed for levels of tryptase, a possible marker of inflammatory processes. At the time of preparing this technical report the laboratory assays on these metabolic markers is not complete and cannot be reported here.

Participants

For the purposes of enlistment of a group of 'responders' and 'non-responders', the response status of all 8-9 year old boys (n=75) who participated in Stage One of the study was determined from their behavioural response to challenge in Mix B and Placebo periods. Using the child's GHA scores for these periods, a GHA difference score was calculated by subtracting the child's Placebo score from their Mix B score. Higher difference scores therefore reflected a more negative behavioural response to challenge by Mix B. The GHA scores were then ranked and children with scores on or above the 75th %ile were classed as responders and those on or below the 25th %ile score as non-responders.

Parents of boys with behaviour scores falling within these groups were contacted by letter and subsequently by telephone and a home visit arranged when they were provided with further details of this stage of the study. Children were invited to participate until 15 boy ('Responders') with scores falling in the upper range and 15 boys with scores falling in the lower range ('Non-responders') were enlisted to the Acute Challenge. Formal consent was obtained from both parent and child.

Methods

The Acute Challenge was conducted over a period of 2.5 hours on the first visit to the School of Psychology, Southampton, and again on the second visit approximately 6 to 7 days later (1 child 5 days later, 1 child = 4 days later). Boys were randomly allocated an order to receive Placebo and Mix B over visits. The amount of additives in Mix B was equivalent to the daily total administered to such children in Stage One of the study. Parents of children were asked to revert to the reduced additive diet followed in Stage One of the study for a period of 24 hours prior to the Acute Challenge. Children and parents were collected by taxi from school or home and returned at the end of the Acute Challenge session. The challenge was administered by capsule in a 15 minute refreshment (fruit and water) break approximately 50 minutes after arrival at the School and parent and child returned home approximately one hour later. Child and parent were provided with £7.50 each in order to buy lunch or other refreshments after completion of the session.

Immediately prior to their visit to the School for the Acute Challenge, parents were asked to collect a (pre-challenge) urine sample from their child and to refrain from giving the children anything to eat in the one hour period prior to the collection of the sample and their visit. Two further urine samples were collected at approximately 15 minutes and 50 minutes post-challenge. In the pre-challenge period, two saliva samples were collected from the child at approximately 50 minutes and 15 minutes pre-challenge. Two further samples were collected at approximately 15 minutes and 50 minutes post-challenge.

After collection of the samples and completion of behavioural assessments in both the pre- and post-challenge periods, the child had short 10 minute breaks to view cartoons.

Behavioural assessment

Conners' Continuous Performance Test II (CPTII) (Conners, 1994)

The children's behaviour in terms of response inhibition and attention was measured using the Conners' Continuous Performance Test II (CPT) (Conners, 1994) employed in Stage One (see Section 7 above). The CPT was administered one two occasions: approximately 30 minutes pre- and post-challenge.

Hillside Behaviour Rating Scale (HBRS)

Their behaviour while completing the CPT was also monitored using the Hillside Behaviour Rating Scale (HBRS) (Gittelman & Klein, 1985). This measure has been used to rate behaviour during testing of children. The HBRS is a seven item scale designed for observational rating of a child's behaviour. Three items assess symptoms of hyperactive behaviour: motor activity, distractibility and impulse control and the remaining 4 items assess more general disruptive behaviour: frustration tolerance, co-operation, interest in tasks and attention seeking. The HBRS provides specific operational definitions of behaviour within each item or domain so that the observer can allocate a score which best describes the child's behaviour during the observation period. In rating of behaviour during testing it has been shown to have high internal consistency (.93) and acceptable levels of interrater relatiabililty (.58 to .76) and to provide a unique source of information to supplement other reports of behaviour.

Data analysis

Scores from the CPT and observational ratings from the HBRS across the pre-challenge and post challenge periods were converted to z-scores and aggregated to produce a Hyperactivity Index (HI) score for the pre-challenge period and a similar score for the post challenge period. Mixed Model (MM) methods were used to analyse the data. In order to reduce the number of parameters in the MM analysis of behaviour the pre-challenge HI score was subtracted from the post-challenge HI score in order to produce a HI difference score. Higher difference scores reflected more negative or challenging behaviour in response to Mix B challenge. Only the results of the behavioural analyses are presented here. The results relating to the role of metabolic factors will be presented at a later date.

Data

Two of the 15 non-responder children failed to return for their second visit for reasons related to swallowing the capsule. The scores for their first visits (1 placebo, 1 active) were included in the MM analysis. The results of analysis are shown in Table 23.1 and indicate a more general trend towards increased behaviour in response to Mix B compared to Placebo and particularly by the group of ‘responders’ compared to ‘non-responders’ where this effect approaches significance (p=.072).

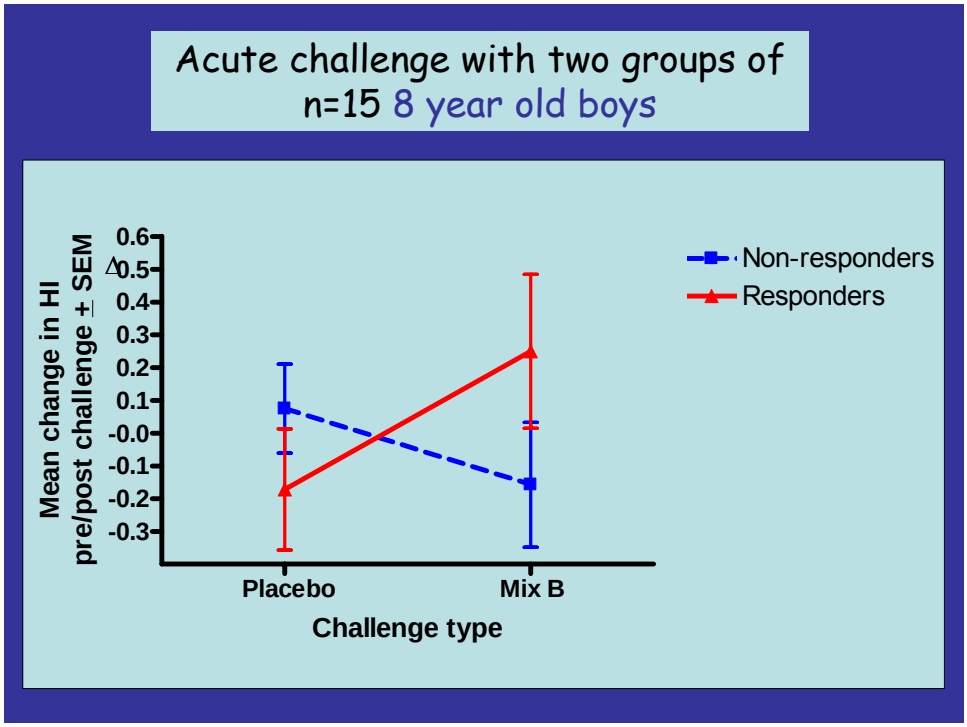
Table 23.1: MM Analyses of effect of Challenge, Responder Group and Challenge x Responder Group on GHA for 8-9yo boys (n=30) participating in Acute Challenge

Mixed Model (n=30) -2LL = 129.23	Factor level	n=	GHA unadjusted mean (SD)	CI	GHA Est. Marg. Mean (SE)	CI	Comparison main effects	Parameter estimate / effect size	CI	P
Challenge F(1, 28.10)=.269 p=.608	placebo	29	-.05 (.63)	-.29 to .19	-.05 (.135)	-.32 to .22	active –v- placebo	-.42	-.92 to .08	.096
	active	29	.05 (.83)	-.26 to .37	.04 (.135)	-.23 to .32				
Responder Gp F(1, 28.39)=.159 p=.693	non-responder	28	-.04 (.62)	-.28 to .20	-.04 (.148)	-.35 to .26	responder –v- non-responder	-.41	-.96 to .13	.134
	responder	30	.04 (.83)	-.27 to .35	.04 (.143)	-.26 to .33				
Challenge x Non-responder F(1, 28.10)=3.50 p=.072	placebo	14	.07 (.51)	-.22 to .37	.08 (.195)	-.32 to .47	)	Non-responder –v- responder Placebo –v- Active: .66 (-.06 to 1.38) p=.072		
	active	14	-.16 (.71)	-.57 to .25	-.16 (.195)	-.55 to .23	)			
Challenge x Responder	placebo	15	-.17 (.72)	-.57 to .22	-.17 (.188)	-.55 to .21	)			
	active	15	.25 (.91)	-.25 to .75	.25 (.188)	-.13 to .63	)			

When the analysis was repeated excluding the 2 children who had missing scores, the pattern of results remained the same with the Challenge x Responder interaction again approaching significance (.70, [-.05 to 1.45], p=.065).

The general trend towards increased behavioural response to Mix B by responders is illustrated in Figure 23.1. The results to date in relation to this Acute Challenge study indicate that additives in food may influence hyperactive behaviour in children.

Fig. 23.1: Hyperactivity index score for responses to Placebo and Challenge by Non-responders and Responders



24. References

Abikoff H, Gittelman R. (1985). Classroom Observation Code: A modification of the Stony Brook code. *Psychopharmacology Bulletin, 21*, 901-909.

Aibkoff H, Gittelman R, Klein DF. (1980). Classroom Observation Code for hyperactive children: a replication of validity. *Journal of Consulting and Clinical Psychology, 48, 555-565.*

Abikoff H, Gittelman-Klein, R, Klein DF (1977). Validation of a Classroom Observation Code for hyperactive children. *Journal of Consulting and Clinical Psychology, 45, 772-783.*

Bateman B, Warner JO, Hutchinson E, Dean T, Rowlandson P, Gant C, Grundy J, Fitzgerald C, Stevenson J. (2004). The effects of a double blind, placebo controlled, artificial food colourings and benzoate preservative challenge on hyperactivity in a general population sample of preschool children. *Archives of Disease in Childhood, 89*, 506-11.

Conners CK, Epstein JN, Angold A, Klaric J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology, 31(5),* 555-562.

Conners CK. (1994). *The Conners continuous performance test.* Toronto, Canada: Multi-Health Systems.

Cook RJ, Zeng L, Yi GY. (2004). Marginal analysis of incomplete longitudinal binary data: A cautionary note on LOCF imputation. *Biometrics, 60,* 820-828.

DuPaul GJ, Anastopoulos AD, Power TJ, Murphy K, Barkley RA. (1994). *ADHD Rating Scale-IV,* Unpublished rating scale.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. (1998). ADHD Rating Scale IV: checklists, norms and clinical interpretation. New York: Guilford, 1998.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey K, Ikeda M. (1997). Teacher ratings of ADHD symptoms: Factor structure and normative data. *Psychological Assessment, 9,* 436-444.

DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, Ikeda M, McGoey K,. (1998). Parent ratings of Attention-Deficity/Hyperactivity Disorder Symptoms: factor structure and normative data. *Journal of Psychopathology and Behavioural Assessment , 20(1), 83-102.*

Epstein N, Erkanli A, Conners CK, Klaric J, Costello JE, Angold A. (2003) Relations between continuous performance test performance meaures and ADHD behaviours. *Journal of Abnormal Child Psychology, 31(5), 543-554.*

Gueorguieva R, Krystal JH. (2004). Move over ANOVA: Progress in analysing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry, 61,* 310-317.

Mallinckrodt CH, Sanger TM, Dubé S, DeBrota DJ, Molenberghs G, Carroll RJ, Potter WZ, Tollefson GD. (2003). Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. *Biological Psychiatry, 53,* 754-760.

Mallinckrodt et al. (2004). Choice of the primary analysis in longitudinal clinical trials. *Pharmaceutical Statistics, 3,* 161-169.

Molenberghs et al. (2004). Analysing incomplete longitudinal clinical trial data (with discussion). *Biostatistics, 5,* 445-464.

Power TJ, Doherty BJ, Panichelli-Mindel SM, Karustis JL, Eiraldi RB, Anastopoulos AD, DuPaul GJ. (1998) The predictive validity of parent and teacher reports of ADHD symptoms. *Journal of Psychopathology and Behavioural Assessment, 20, 57-81.*

Zhang S, Faries DE, Vowles M, Michelson D. (2005). ADHD Rating Scale IV: psychometric properties from a multinational study as a clinician-administered instrument. *International Journal of Methods in Psychiatric Research, 14(4), 186-201.*