Research Report DFE-RR006

UK Resilience programme evaluation:

Second interim report

Amy Challen, Philip Noden and Anne West (1) and Stephen Machin (2)
London School of Economics and Political Science (1)
London School of Economics and Political Science/University College London (2)

Department for Education

This research report was written before the new UK Government took office on 11 May 2010. As a result the content may not reflect current Government policy and may make reference to the Department for Children, Schools and Families (DCSF) which has now been replaced by the Department for Education (DFE).

The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

Evaluation of the UK Resilience Programme

There are increasing concerns about children's well-being in the UK, their behaviour, and the low academic attainment of a large fraction of the population.[1] Recently, the potential and duty of schools to promote pupils' well-being has been stressed as part of the Every Child Matters agenda. In September 2007, three local authorities (South Tyneside, Manchester and Hertfordshire) piloted a programme with Year 7 pupils in 22 of their schools, with the aim of building pupils' resilience and promoting well-being: the UK Resilience Programme. More schools have since begun teaching the programme.

The evaluation aims to investigate whether the programme (previously trialled in small samples) can be delivered at scale; whether it has an impact on children's well-being; and if so, whether this will have an impact on behaviour, attendance and academic attainment.

The first interim report was published in April 2009 and gives an overview of the UK Resilience Programme and its implementation, describes the evaluation, and offers preliminary findings about programme impact. The report also contains a bibliography and descriptions of previous research on the Penn Resiliency Program (the curriculum on which UKRP is based), and describes the curriculum in detail. In this second report we frequently refer to the first report[2] for background information and previous findings, and it can be found online at:

http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RR094%20(1).pdf

while the corresponding four-page research brief can be found at:

http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RB094.pdf

The qualitative work for this report was carried out by Dr. Philip Noden and Prof. Anne West.[3]

The quantitative work is by Amy Challen[4] and Prof. Stephen Machin[5].

---

[1] See, for example, the recent UNICEF report "An overview of child well-being in rich countries" which puts the UK at the bottom of a list of 21 advanced countries: http://www.unicef-irc.org/publications/pdf/rc7_eng.pdf

[2] Referred to as Challen et al. (2009). The first and second interim reports are by the same authors.

[3] Both Education Research Group, Department of Social Policy, London School of Economics and Political Science.

[4] Centre for the Economics of Education and Centre for Economic Performance, London School of Economics and Political Science.

[5] Centre for the Economics of Education and Centre for Economic Performance, London School of Economics and Political Science, and Department of Economics, University College London.

CONTENTS

1. Executive Summary 4
2. Policy and Delivery Implications 5
3. Quantitative Findings 7
    - Summary of Quantitative Findings 7
    - Data collected 8
    - Explanation of measures used 13
    - Treatment and control groups 14
    - Programme impact on symptoms of depression 17
    - Programme impact on symptoms of anxiety 25
    - Programme impact on absence from school 26
    - Conclusion 28
    - Implementation and policy issues 29
    - Tables 32
4. Qualitative Findings 56
    - Summary of Qualitative Findings 56
    - The case study schools in 2009 57
    - A drift to non-teachers? 62
    - The importance of senior management backing 67
    - Finding a place in the curriculum 70
    - Interviewees' view on the UKRP 74
    - Conclusion 81
5. Appendix 84
6. References for Empirical Evaluations of PRP and UKRP 85

1. Executive Summary

Evaluation of the UK Resilience Programme, 2nd Interim Report

The UK Resilience Programme (UKRP) aims to improve children's psychological well-being by building resilience and promoting accurate thinking. Three local authorities launched it in the academic year 2007-08, with workshops delivered to Year 7 pupils in secondary schools. This report presents interim findings from the UKRP evaluation, commissioned by the Department for Children, Schools and Families.

Methodology

Information on pupils' well-being was collected through questionnaires administered before and after the programme to pupils who had participated in the first year of UKRP workshops and to a control group, as well as at several follow-up points. In addition, interviews with pupils, facilitators (workshop leaders) and school managers were carried out in 10 of the 22 secondary schools involved in the programme at the end of the academic year 2007-08, and follow-up interviews were carried out in 9 of these 10 schools in the autumn term of 2009-10. The quantitative work examines what has happened to the original cohort of pupils over time (those in workshops in 2007-8), while the interviews look at the current implementation of UKRP in the original case study schools.

Key findings

- The quantitative work found a significant short-run improvement in pupils' depression symptom scores and school attendance rates. There was also an impact on anxiety, but this was smaller, and concentrated in a few groups of pupils: boys, particularly boys with SEN or FSM entitlement, and lower-attaining girls.
- The size of the impact varied by how workshops were organised. Weekly workshops showed a larger impact than those timetabled fortnightly.
- The impact also varied by pupil characteristics, and was larger for pupils with Special Educational Needs (when the outcome was anxiety or depression); for pupils entitled to free school meals (anxiety and attendance); for pupils who had not attained the national target levels in Key Stage 2 exams (depression, anxiety and attendance); and for pupils with worse initial scores for symptoms of depression or anxiety (depression and anxiety).
- On average the effect of the workshops lasted only as long as the academic year, and had faded by the one-year follow-up questionnaire in June 2009. However, there was still an impact for certain groups at follow-up, particularly for pupils who had not attained the national target levels at Key Stage 2 in English or maths.
- Return visits to nine of the case study schools revealed that seven of the nine schools were continuing to deliver the UKRP to all Year 7 pupils.
- Facilitators were extremely positive about the ideas underlying the programme and about the training they had received. Most reported that they used the skills themselves.
- Facilitators found the curriculum materials didactic and thought they could be improved. Many felt that some pupils struggled with the programme content and materials.
- Pupils were generally positive about the programme. Interviews for the First Interim Report suggested that pupils had applied UKRP skills in real life situations, and some interviewees showed a good understanding of elements of the programme.
- Future quantitative analysis will examine the longer-run impact on attendance, academic attainment and psychological well-being.
- The final report will be available in early 2011.

2. Policy and Delivery Implications

Here we list potential policy implications of the results presented in this report and the findings of last year's report. Many of these points are considerations rather than recommendations, but they do highlight issues around implementing the programme.

1) It is essential that the programme has the backing of senior management within schools.

2) A preferred model of delivery for the UKRP, based on the recommendations of the course developers and the findings of this study, might involve 18 weekly sessions delivered to groups of no more than 15 pupils. Only two of the nine case study schools were delivering the UKRP in accordance with this preferred delivery model. It is clear there are pressures on the UKRP within schools, and these may arise from financial or timetabling demands or from pressure to improve pupils' levels of attainment.

3) There is a drift evident in some schools towards the programme being delivered by non-teachers, in part because of the pressures noted above. Such a drift may reduce the pool from which potential trainees may be selected.

4) It is important that school staff delivering the programme know how the school's child protection arrangements work, and are aware that the programme may lead to disclosure of serious problems by pupils. Staff need to be adequately prepared for and supported throughout the programme in order to deal with these issues.

5) The UKRP was intended to be a universal programme, but some schools have chosen to target pupils for inclusion in workshops. It is not clear which model is preferable, and this will probably depend on the situation of each school. However, the following points are worth bearing in mind:

    - Based on the quantitative analysis, certain groups of pupils appeared to benefit more from the workshops, particularly those who did not achieve the national targets at Key Stage 2, pupils with SEN, and pupils who started the school year with higher levels of depression or anxiety symptoms.

    - However, the measured impact on these pupils is the impact of the programme delivered to 'universal' or mixed workshop groups, not of groups consisting entirely of targeted pupils. One cannot therefore assume that the same impact would be obtained if workshop groups were targeted.

    - Some schools that did run workshops entirely with targeted pupils reported these as being very difficult to manage and not very successful compared to more mixed groups.

    - The same applies to levels of academic attainment: many facilitators commented that SEN groups or lower set groups did not go well, or that the presence of more able or more literate pupils aided the success of the lessons.

    - Although facilitators and other school staff often appeared to assume that higher ability pupils were naturally more resilient, or had fewer problems, almost all facilitators claimed to use the UKRP skills themselves. It therefore seems unlikely that higher ability pupils or those with better psychological well-being would be unable to benefit from the skills.

- Even if pupils were to be targeted for inclusion in workshops, it is important that they should be targeted appropriately. Previous research suggests that school staff tend to identify pupils with behaviour problems rather than those with emotional difficulties, yet the programme is primarily designed to address the latter. The process of targeting would also need to be carefully considered.

- Participation in programmes perceived to be targeted and remedial can attract stigma for those who participate. Universal programmes avoid this.

- The measures used in the evaluation are sensitive to differences in the severity of symptoms of depression and anxiety, but are not good at distinguishing between children who have few or no symptoms. For instance, they would not be able to detect any improvements in well-being for pupils who showed no initial symptoms of depression, although this would not necessarily mean that these children did not benefit.

- The skills pupils used most (as reported by both pupils and facilitators) were the interpersonal skills around negotiation and assertiveness, and techniques for self-control. Since all pupils are likely to experience conflict and problems around everyday social interactions it is likely that all pupils could benefit from the workshops, at least in these areas.

3. Quantitative Findings

Summary of Quantitative Findings

The first UKRP workshops were delivered to Year 7 pupils in 22 participating schools in 2007-8. In the First Interim Report (Challen et al., 2009) we provided an assessment of the short-run impact of the workshops (to July 2008), finding that on average they had a small but significant impact on pupils' depression and anxiety scores. We also found differences in the size of the effect of the programme based on the timing and frequency of the workshops (weekly workshops starting at the beginning of the academic year had more of an impact), and by pupil characteristics (lower attaining and more disadvantaged pupils gained more, as well as pupils who started the year with worse symptoms).

In this follow-up quantitative analysis we look at the same cohort of pupils and examine the impact of the programme at the one-year follow-up point in June 2009, comparing this to the impact seen in the first year of the workshops. We also improve upon the method of analysis we used in the 2009 report. We use two samples of pupils: the full sample of all UKRP and control pupils who were in Year 7 in 2007-8, and a subsample of these in which programme and control pupils are well matched on a variety of characteristics. We use both in our analyses, and obtain similar results.

We find an average improvement in pupils' depression symptom scores and school attendance as a result of the workshops, although this has faded by one-year follow-up for the depression score (we have not yet been able to examine this for attendance as the data are not yet available). There was also an impact on anxiety, but this was smaller than the impact on depression, and concentrated in a few groups of pupils: boys, particularly boys with SEN or FSM entitlement, and lower-attaining girls. As we found in the 2009 report, we find important differences in the effects of the workshops in terms of how they were organised, with weekly workshops starting at the beginning of the academic year showing more impact.

The impact of the programme also varies by pupil characteristics: in general, lower attaining and more disadvantaged pupils appear to gain more from the workshops (as we found in the previous report), and in some cases the programme impact has not faded for these groups by the one-year follow-up. The average improvement in absence rates appears to be similar across different ways of organising workshops, and is also similar for most groups of pupils. It is equivalent to an improvement of about 1.8 more school days attended over the course of the year.

Further study of the longer-run results will be possible once the next wave of data has been collected in June 2010.

Quantitative analysis

Introduction

For the purposes of comparison, we are providing similar analyses as for the 2009 report. However, we now include the most recent follow-up data, consider a larger sample of pupils, and evaluate the effect on an additional outcome (absence from school). All pupils who were in Year 7 in 2007-08 are included in the analysis below if they completed enough questions from the questionnaires or have sufficient records available from the National Pupil Database (for absence data).

However, not all implementations of the UKRP intended for inclusion in the evaluation appear to have produced comparable intervention and control groups. Having similar workshop and control groups is important: unless we are sure that the pupils included in the workshops and those in the control group were similar to begin with, we cannot be sure whether any differences between the two groups at the end of the workshops were due to the effects of the programme, or were simply due to pre-existing differences between them. In the analysis we therefore present results for both the full sample, and for the group (based on workshop timing) that has well-balanced intervention and control groups.

Data collected

Year 7 pupils were surveyed at the beginning of the school year before the start of the intervention, at the end of the intervention and at the end of the academic year (two or three times in one year, depending on workshop timing) – see Figure 1: Timeline of workshops and data collections.[6] There was a further follow-up survey in July 2009, and there will be another in July 2010.

Tables 1a and 1b give summary statistics for the questionnaires completed in the evaluation so far. The 'Year 7 cohort' refers to pupils who were in Year 7 in 2007-08, both workshop and control pupils. 'Year 8 cohort' refers to pupils who were in the year above these pupils in some schools and are an additional control group. In the analysis that follows we use only the Year 7 cohort, but the number of questionnaires collected from the Year 8 cohort is included in Tables 1a and 1B below.[7]

Because this is a panel dataset pupils will usually appear more than once, except in cases of attrition (pupils leaving the sample), or addition to the sample (e.g. new pupils arriving at the school after the baseline). These multiple observations of (most) pupils give us a difference in the tables below between the number of questionnaires collected by wave (Table 1a), and the number of pupils represented at each data collection (Table 1b: 'represented' means that we have at least one questionnaire for them, either a pupil or a teacher one).

[6] Only 9 of 22 schools were involved in the mid-year data collection (Wave C, around February 2008), as this only concerned schools that were starting or finishing a set of workshops around this time.

[7] The first data collection involving Year 8 pupils occurred in July 2007; this is not shown on the timeline.

Figure 1: Timeline of workshops and data collections

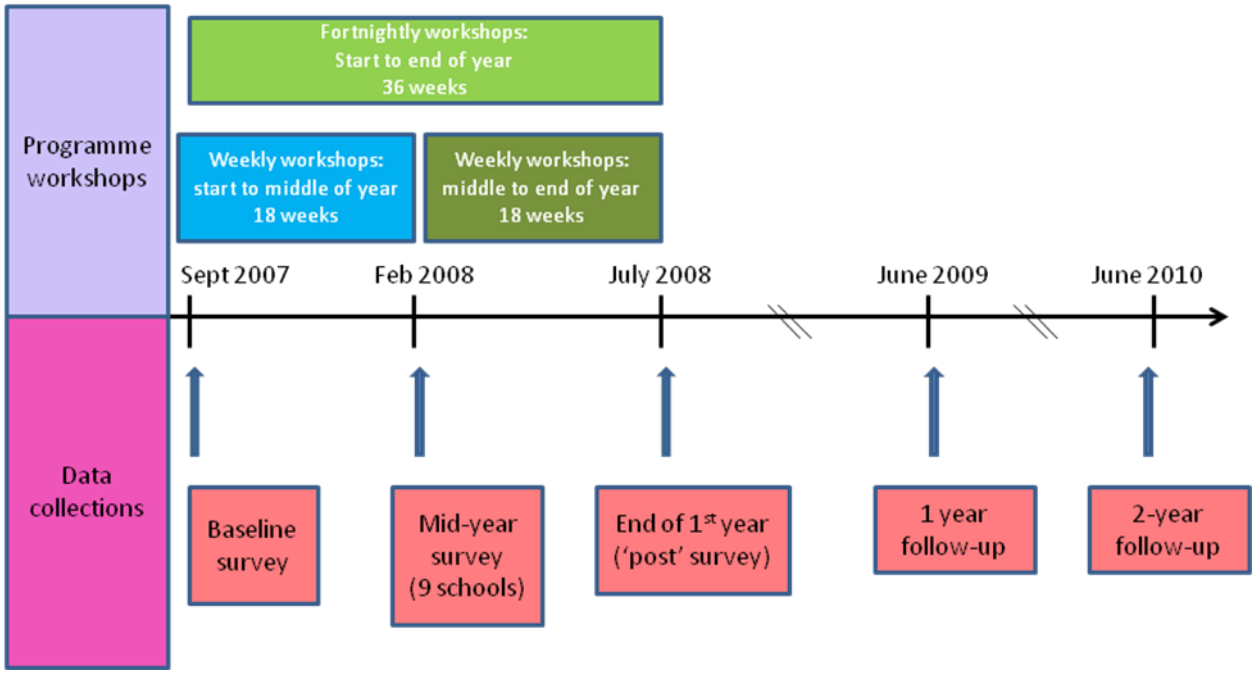

Table 1c presents more information on attrition (or incomplete panels of data) for the Year 7 cohort only. The columns are headed by the number of times a pupil is represented in our sample: this ranges from once (we have information on the pupil for only one point in time), to four times (we have information on them for each of the four data collections involving the Year 7 cohort, including the February 2008 data collection). The second last row total (pupils*waves) tells us how many pupil-wave observations we have. The last row gives the number of pupils in each of these groups. Since not all schools were involved in the February 2008 data collection (Wave C), most pupils will not have been surveyed at this point, so a pupil with no missing data points could appear either three or four times. Note that the proportion of pupils for whom we have either three or four data points is high: 3715, which is 90% of the 4108 Year 7 pupils who appear in our data at least once. Only 229 (6%) of pupils appear only once, and of these, only 86 (2% of the total cohort) appear in Waves B to D, implying a very low attrition rate. The 143 pupils who appear once but who only appear in Wave E are additions to our baseline sample.[8]

Tables 2a, 2b, 3a and 3b present similar data on the analysis samples – these are the samples of pupils we use in the quantitative analysis below. Of the possible 4108 pupils from this cohort who have at least one questionnaire in the evaluation, we use about 3150 pupils for the evaluation sample: this consists of all pupils in the cohort with enough observations at different waves for us to undertake the desired analysis. To be included these pupils had to

[8] Note that pupils who first appeared in Wave E are not necessarily new to the school; they may have been absent at previous data collections, or have opted not to fill in the questionnaires.

have measures for at least waves B and E, and for C or D as well to be included in the evaluation of programme impact at these points.

The first sample we use is the full sample of the control group plus all pooled workshop groups (there are three workshop groups, defined by workshop timing – see 'Treatment and Control groups' section below). This gives us a reasonably large sample with a roughly equal split between workshop (1607) and control group (1546) pupils for the psychological outcomes for which we needed data from pupil questionnaires (Table 2a). However, we can use a larger number of pupils when our outcome is pupil absence rates. These data come from the National Pupil Database, and are available for pupils in our sample for as long as they attend state schools in England, so there is an even lower rate of attrition. Table 2b shows that when we restrict the sample to pupils with absence data for both academic years (2006-7 and 2007-8), we obtain a sample of 3,810 pupils (1901 workshop; 1909 control), and 7,620 observations.

We use the second, more restricted sample in addition to the full sample because here treatment and control groups appear more similar at baseline. This means we can be more confident that post-workshop differences between the two groups will be driven by the exposure to the workshops, rather than any pre-existing differences in pupil characteristics. Table 3a gives the number of observations we can use when using psychological variables as outcomes: the control group is the same as before (1607 pupils), but the treatment group consists only of pupils who took part in Group 1 workshops: those that were held weekly and started at the beginning of the academic year. This leaves us with 373 workshop pupils, and a total sample of 1980 pupils (6826 observations). Table 3b gives the sample size for the Group 1 workshop sample when the outcome is the absence rate. Here we have a total sample of 2355 pupils (4,710 observations).

Table 1a: Number of questionnaires returned throughout evaluation

		Pupil questionnaires returned		**Teacher questionnaires returned**		**Total**
Data collection	**Dates**	**Year 7 cohort**	**Year 8 cohort**	**Year 7 cohort**	**Year 8 cohort**	
Wave A	July 2007	0	2,152	0	1,887	**4,039**
Wave B	September 2007	3,642	0	3,184	0	**6,826**
Wave C	February 2008	1,786	0	1,757	0	**3,543**
Wave D	July 2008	3,638	2,113	3,508	1,792	**11,051**
Wave E	June 2009	3,621	2072	3,544	1975	**11,212**
Total	**July 2007-Sept 2009**	**12,687**	**6,337**	**11,993**	**5,654**	**36,671**

Table 1b: Number of pupils represented at each data collection

		Number of pupils with at least one questionnaire (pupil or teacher)		Total
Data collection	**Dates**	**Year 7 cohort**	**Year 8 cohort**	
Wave A	July 2007	0	2,152	**2,152**
Wave B	September 2007	3,784	0	**3,784**
Wave C	February 2008	1,943	0	**1,943**
Wave D	July 2008	3,814	2,195	**6,009**
Wave E	June 2009	3,856	2,213	**6,069**
Total	**Sept 2007-June 2009**	**13,397**	**6,560**	**19,957**
Total pupils		**4,108**	**2,370**	**6,478**

Table 1c: Attrition in Year 7 (workshop) cohort

		Number of times a pupil appeared in sample (pupil or teacher questionnaire)				
Data collection	**Dates**	**1**	**2**	**3**	**4**	**Total**
Wave B	September 2007	70	85	1,934	1,695	**3,784**
Wave C	February 2008	3	48	197	1,695	**1,943**
Wave D	July 2008	13	120	1,986	1,695	**3,814**
Wave E	June 2009	143	75	1,943	1,695	**3,856**
Total (pupils*waves)	**Sept 2007-June 2009**	**229**	**328**	**6,060**	**6,780**	**13,397**
Total respondents		**229**	**164**	**2,020**	**1,695**	**4,108**

Analysis samples

Table 2a: Full sample - pooled treatment groups and control group

Number of observations for use in analysis where psychological variables are the outcome

Data collection	Dates	Combined workshop groups	Control group	Total
Wave B	September 2007	1,607	1,546	**3,153**
Wave C	February 2008	564	826	**1,390**
Wave D	July 2008	1,552	1,447	**2,999**
Wave E	June 2009	1,607	1,546	**3,153**
Total questionnaires	**Sept 2007-June 2009**	**5,330**	**5,365**	**10,695**
Total respondents		**1,546**	**1,607**	**3,153**

Table 2b: Full sample: absence measure, pooled treatment groups and control group

Number of observations for use in analysis where absence rate is the outcome

Academic year	Combined workshop groups	Control group	Total
2006-7	1,901	1,909	**3,810**
2007-8	1,901	1,909	**3,810**
Total questionnaires	**3,802**	**3,818**	**7,620**
Total respondents	**1,901**	**1,909**	**3,810**

Table 3a: Group 1 workshop group and control group

Number of observations for use in analysis where psychological variables are the outcome

Data collection	Dates	Group 1 workshop pupils	Control group	Total
Wave B	September 2007	373	1,546	**1,919**
Wave C	February 2008	353	826	**1,179**
Wave D	July 2008	362	1,447	**1,809**
Wave E	June 2009	373	1,546	**1,919**
Total questionnaires	**Sept 2007-June 2009**	**1,461**	**5,365**	**6,826**
Total respondents		**373**	**1,607**	**1,980**

Table 3b: Group 1 workshop group and control group: absence measure

Number of observations for use in analysis where absence rate is the outcome

Year	Group 1 workshop pupils	Control group	Total
2006-7	446	1,909	**2,355**
2007-8	446	1,909	**2,355**
Total questionnaires	**892**	**3,818**	**4,710**
Total respondents	**446**	**1,909**	**2,355**

Explanation of measures

Symptoms of depression are measured using the Children's Depression Inventory (CDI). This scale has 27 items in the original version, but this study omits the item about suicidal ideation and so it contains only 26.[9] Each of these is scored as 0, 1 or 2 depending on the severity of the symptom: 0 indicating no symptoms of depression on that item according to the child's response; and 2 indicating strong symptoms. Item scores are then summed to create a total score between 0 and 52, where higher scores indicate worse symptoms.[10] However, since the scale primarily measures deviations from well-being, rather than degrees of positive well-being, the distribution of the total score is highly skewed. A large number of pupils have very low scores: over 10% score 0 or 1, and over 50% score 7 or lower, and this is true for each of Waves B, C, D and E. We therefore encounter a 'ceiling effect' on depression symptoms scores, as pupils scoring 0 in the baseline at the beginning of the year cannot improve their scores.

Symptoms of anxiety are measured using the Revised Children's Manifest Anxiety Scale (RCMAS). This scale contains 28 items, plus 9 more to form a 'lie scale' used to detect responses motivated by social desirability (though the latter scale is not used in the present analysis). Each item of the main scale asks about whether a symptom of anxiety is typical of the child or not, and is scored as 1 if the response is 'yes' and 0 if 'no', giving a maximum possible summed score of 28 with higher scores indicating worse symptoms. Again, the distribution of scores is highly skewed, with over 10% of pupils scoring 0 or 1 and over 50% scoring 8 or lower.

Behaviour is measured using the self-report and teacher-report versions of the Goodman SDQ. Although no analysis on these scores is reported here, pupils' scores on these measures are used to check the comparability of the treatment and control groups (see Tables 4-7). The SDQ total difficulties score is comprised of 20 items, each scored 0, 1 or 2 according to the perceived severity of the symptom. This gives a minimum possible score of 0 and a maximum of 40, with higher scores indicating more (and more severe) symptoms.[11] The distribution of scores is highly skewed, particularly for the teacher version: 50% of all pupils score 5 or lower on the teacher SDQ; and more than 50% score lower than 11 on the pupil version.

Life satisfaction is measured using the Huebner Brief Multidimensional Students' Life Satisfaction Scale, which has five items asking about satisfaction with particular domains of a child's life and one asking about overall life satisfaction.[12] This is scored on a 7-point scale, giving a minimum possible summed score of 6 and a maximum of 42, with higher scores indicating greater life satisfaction. Again the distribution is skewed, with over 50% of pupils scoring 35 or more. Although no analysis on these scores is reported here, the baseline

[9] The item on suicidal ideation is often omitted when using the inventory in universal (as opposed to psychiatric) populations. The item was not deemed appropriate or necessary for use in schools, so has been omitted at all stages of this study.

[10] If more than 10% of items are unanswered then the assessment is considered invalid. When up to 10% of items are missing these scores can be replaced by the mean of the non-missing items in order to create a total score.

[11] The assessment is valid if at least 3 items of each of the four difficulties subscales have been completed.

[12] The domains are: family, friends, school, oneself, and where the respondent lives.

values of this measure are reported in Tables 4-7 in testing for the similarity of the treatment and control groups. Note that on this scale higher scores indicate greater well-being, unlike the other four scales for which the reverse is true.

Annual absence from school is measured as the fraction of school sessions for which pupils were absent during the academic years 2006-7 and 2007-8. This is the sum total of authorised and unauthorised absences during the year.[13] The data on pupils' gender, SEN status, entitlement to free school meals, Key Stage 2 attainment, and absence are obtained from the National Pupil Database/Pupil Level Annual Schools Census (NPD/PLASC).

Treatment and control groups

Tables 4-7 present the treatment and control group means for 11 variables for the full sample and for the three different experiments separately, including the p-values from mean-comparison tests. These tests indicate how different the two samples are at the baseline (September 2007), and therefore tell us how well matched the workshop and control groups were before any of the workshops began. They provide an estimate of how likely it is that an equally large difference between the two means could have arisen by chance if the samples were in fact similar. The first six variables in the upper panel of each table are potential outcome variables. The first five are psychological and behavioural variables obtained from the questionnaires we administered in September 2007; the sixth is annual absence, obtained from NPD/PLASC for the academic year 2006-07, i.e. the year before the workshops and the year before these pupils transferred to secondary school. The five variables in the lower panels are socio-demographic and educational attainment variables obtained from the NPD and matched into our dataset. The data on Key Stage 2 attainment comes from the KS2 attainment tables in the NPD for summer 2007. The values from the other variables are taken from PLASC for the academic year 2007-08, unless this record is missing, in which case values from the previous year (2006-07) or the following year (2008-09) are used. In this way we minimise the number of pupils for whom we lack demographic information.

Table 4 presents statistics for the full Year 7 cohort when they were surveyed at baseline. This includes all pupils who were in Year 7 in 2007-8, in either workshop or control groups, and who completed enough questions and questionnaires throughout the course of the data collections to be included in the evaluation sample (so this does not include the year-above control groups, who do not have a baseline measure). It is clear that there are significant differences between the workshop and control groups: the p-values of tests of equality of means between the treatment and control groups are less than 0.1 for 7 out of 11 variables, suggesting that there is (at most) a 10% chance that the observed differences between these groups would have arisen by chance if they were in fact similar. (The smaller the p-value, the less likely that such a difference could have arisen by chance.) Importantly, there are significant differences between treatment and control at baseline for both the depression score (p=0.03) and the anxiety score (p=0.08), which are the first two of the three outcome

[13] We use this measure of absence because the distinction between authorised and unauthorised absence is not reliably recorded in the data, but the total number of absences is usually quite accurate.

variables we will consider in our analysis below. The two groups do seem similar in terms of the mean absence rate from school in the previous year (p=0.74). However, the groups also differ significantly in terms of demographic and attainment characteristics (the five variables in the bottom panel of Table 4). Although we cannot reject that they are have a similar composition in terms of gender (p=0.18), the full treatment group and the control group are significantly different at the 5% level for the others: on average the treatment group as a whole has lower levels of SEN, FSM eligibility, and higher prior attainment than the control group.

Since we have data on these characteristics, we can control for them in our analysis, and thereby control for any differences between treatment and control groups that are not due to the causal impact of workshop participation. However, given that there are these observable differences between the two groups, it is likely that there are also differences between them that we cannot observe (we do not have data on them), and therefore cannot control for. If these unobserved differences are correlated with assignment to the treatment or control group, then if we use poorly matched groups we cannot be sure that differences in outcomes between treatment and control are due to the programme and not due to other factors. This is why it is important to have well-matched treatment and control groups, and why we will not only use the full sample in our analysis.

The following three tables present the same information as Table 4 but for each configuration of workshops separately. Table 5 presents descriptive statistics for pupils who were in 'Group 1' workshops, relative to the control group (the control group is the same as for the full sample, and for the other two groups of workshops discussed below). These are workshops that took place at least once a week from the beginning of the school year, and finished around the middle of the school year (roughly September 2007 to February 2008). This is a similar sample to the one that was used for most of the analysis in the First Interim Report, but we have included more observations this year.[14]

The treatment and control groups are well matched: it is clear from Table 5 that most of the p-values from the mean comparison tests are well above 0.1. In fact, the only two variables presented for which there are statistically significant differences in means are Key Stage 2 maths attainment and absence rate in the previous year. On average, pupils in the treatment group had significantly higher KS2 maths scores than those in the control group. The difference between the two groups is not large (0.13 of a level), but it is interesting that on average pupils in the control group did not attain the national target level 4 in Key Stage 2 maths, implying that they have somewhat below average attainment. This difference could have resulted from the way schools assigned class groups to UKRP workshops: if the assigned classes were setted, for instance in science or English sets, the treatment group's academic attainment might differ from the school average. This was not necessarily deliberate, as many schools simply had to choose classes to participate in UKRP workshops based on which ones fitted the timetable. Perhaps more worrying is the difference in

[14] In the First Interim Report (Challen et al. 2009) we only included pupils in schools which had a Year 7 control group, as we were concerned about the possibility of differences between schools driving the results. Here we include all Year 7 pupils in the relevant categories for whom we have sufficient data and control for the school that a pupil attended in September 2007, which produces similar results to the estimates we presented last year but allows us to use a larger sample.

absence rates, as it is one of the outcome variables we will be examining. Pupils in these workshops had a higher absence rate in their last year of primary school than pupils in the control group. We suspect that this was largely due to pupils in these two groups coming from different schools in different proportions, meaning that they are likely to have different average levels of attendance and attainment. (Schools are not equally represented in both groups, as the proportion of the yeargroup included in workshops varied greatly by school, ranging from 11% to 100% of the cohort.) That is to say, the differences are due to between-school differences in these variables rather than within-school differences. Nevertheless, in our analysis below our preferred specification uses pupil fixed effects, using multiple observations of the same pupil over time to control for all the characteristics of the pupil which are fixed through time, and this should control for differences in KS2 attainment and prior absence rates. Moreover, we control for the school that each pupil attended in September 2007 and allow for different trends by school, and this should take account of many differences due to schools.[15]

The second treatment group consists of weekly workshops that started mid year and finished near the end of the school year (roughly February to July 2008), and baseline statistics for this sample and the control group are presented in Table 6. Here it is clear that the pupils in the workshops are on average significantly different from the control group pupils on the psychological measures, on one of the behavioural measures, and in terms of their absence rate (the p-values are less than 0.05 for all but one of the six possible outcome variables, meaning that there is less than a 5% chance in each case that such a large difference between the treatment and control groups would have arisen by chance if they were in fact similar). We know that some schools selected pupils for this second batch of workshops based on perceived psychological need, so it is not surprising that they look different to the control group. However, they seem reasonably well matched on the five variables in the lower panel – the socio-demographic and attainment measures.

Table 7 presents measures for the third group of workshops, which took place fortnightly and lasted from the start of the academic year to (near) the end. This treatment group seems well matched to the control group in terms of the psychological scores at baseline, with the exception of life satisfaction (p=0.09, with the control group scoring higher than the treatment group, meaning that on average they report being more contented). However, the difference in the absence rate is significant at 0.1%. Moreover, these pupils are very different in terms of all five socio-demographic and attainment variables: the treatment group is significantly more female, higher attaining, less likely to be entitled to free school meals, and less likely to have special educational needs than the control group. Many of the schools that timetabled workshops in this pattern included all of their Year 7 pupils in workshops and so do not have

[15] There could still be a problem if baseline differences give rise to different trends, and these trends are correlated with workshop/control assignment. For instance, if it were the case that more academically able children become happier at secondary school while less able children become less happy, and a school assigned only its top set maths pupils to workshops, we might attribute the improvement in these pupils' subjective well-being to the workshops when in fact it was due to their academic ability. However, we do not think that the mismatch in maths attainment is a particular problem for us, as we find that pupils who did not attain Level 4 in KS2 maths actually seem to benefit more from the workshops than those who did, see Summary below. Given that the descriptive statistics suggest that there is a lower proportion of low attaining pupils in our treatment groups this mismatch would in fact bias downwards, not upwards, our estimate of the average treatment effect.

a within-year control group; moreover, they are on average less deprived and higher attaining than the sample as a whole. This could make comparisons more difficult, as treatment and control pupils are in different schools, so are probably more different to begin with, and may be subject to different environments and events throughout the course of the year. It is therefore not surprising that treatment and control pupils appear to be significantly different at baseline. As mentioned above, in our analysis we can control for many differences between schools, but with so many obvious differences between the pupils in this treatment group and the control group this is probably not the best sample to rely on to estimate the programme impact.

Thus, because of the significant baseline differences between the treatment and control groups when using the full sample, we propose to conduct and present our analyses on both the full sample and on the first group of workshops (the September to February weekly group of workshops), which is well matched to the control group. We can be reasonably confident of any results we obtain that are robust to using both of these two samples.

Programme impact on symptoms of depression

First we present the raw differences in the mean depression symptoms score between treatment and control groups, without controlling for pupil or school characteristics (Table 8). Then we will present an analysis of the treatment effect using an econometric specification to control for pupil characteristics, which should go some way to correcting any bias produced as a result of mismatched treatment and control groups (Table 9). Finally we will use our preferred specification and examine heterogeneity in treatment effect by pupil characteristics (Table 10). We follow a similar pattern for analyses of the programme impact on the anxiety score and on absence.

Raw scores

Table 8 presents the raw analysis of the programme impact when the outcome is the depression symptoms score (measured by the Children's Depression Inventory - CDI score). This compares with Table 29 of Challen et al. (2009), but here we are using a larger sample of pupils as discussed above, and have added a later follow-up data point, one year later (data collected in June 2009). Here we present the mean depression symptoms score for the three separate treatment groups and the control group at three points in the year, and one follow-up: the start of the academic year in September 2007; the mid-year data collection point (February-April 2008); the end of the year in June-July 2008; and at one-year follow-up in June 2009. As shown in the table, the treatment and control groups are not perfectly matched at the beginning of the year: for all three experiments the treatment group scores higher (worse) than the control group, though only for the mid-to-end year treatment (middle panel) is the gap between the two significant, with a p-value of less than 0.01.[16] [17] Looking at

[16] Remember that for the depression score, a higher score indicates worse symptoms, so a decline in the scores is an improvement.

[17] If the difference between two values is statistically significant this means that it is unlikely that an equally large difference would arise by chance if in fact the two values were similar. For instance, if a difference is significant at the 10% level (p<=0.1), this means that there is only a 10% chance that it could have arisen by chance; if it is significant at the 1% level (p<=0.01), there is only a 1% chance of

the first panel, we see that by the middle of the year, pupils who were in UKRP workshops score lower (better) than the control group. The difference-in-difference coefficient is obtained by subtracting the mid-year mean from the baseline mean for the control and treatment groups, then taking the difference between these two to obtain the overall effect of treatment: it measures how the treatment group has changed relative to how the control group has changed. This shows a significant negative change (coefficient=-1.12, p=0.01) by the middle of the school year, implying that those who were in these workshops improved their depressive symptoms score relative to those in the control group. There is still a significant improvement at the end of the year (July 2008), although one year later (June 2009) this has faded out and is no longer significant (p=0.72).

The second panel concerns the workshop groups that were treated from the middle of the year to the end (Group 2 workshops). They are clearly quite different from the control group at the beginning of the year, on average scoring significantly higher (meaning worse symptoms), and although their mean score declines through the year it remains higher than that of the control group. The difference-in-difference estimates for both the end of the first year and the one-year follow-up are not significant, implying that on average, and without taking into account pupils' characteristics, there is no measured impact on depression symptoms of this group of workshops.

The third panel of Table 8 shows the control group compared to UKRP pupils who participated in Group 3 workshops that lasted all academic year. Here the treatment group starts the year scoring slightly worse than the control group, but the difference is not significant. By the end of the year the treatment group score has deteriorated and the control group has improved slightly, giving a positive difference-in-differences coefficient, i.e. suggesting an adverse effect on well-being of participating in these workshops. This effect persists to the one-year follow-up, and both these estimates are significant at 1%. As mentioned above, since these are raw scores we do not control for differences between pupils or schools in this table, so it could be that this effect is due to pre-existing differences between the treatment and control groups.

Adding controls for pupil and school characteristics

The overall treatment effect for all three groups of workshops combined is presented in Table 9 (compare with Table 28 of the First Interim Report). Because here we use an econometric specification we are able to control for other factors that could affect the estimated programme effect. Columns 1-4 present the short-run effect of the workshops, measured either in the middle of the year immediately after the end of the workshops; at the end of the academic year (immediately post for Groups 2 and 3); or both. Columns 5-8 present the results for the one-year follow-up, with data collected in June-July 2009. The coefficient on 'Treated*PolicyOn' gives the difference-in-differences estimate, although here the outcome variable is the standardised depression symptoms score, so the size of the estimate is not easily comparable to the previous table.[18] The first column only controls for the month in which the questionnaire was filled in (a control for reporting and seasonal

its arising by chance, etc.

[18] Standardising scores involves subtracting the mean score and dividing by the standard deviation to give a standardised score with a mean of 0 and a standard deviation of 1. This does not change the results we obtain, but makes interpretation and comparison of the coefficients easier.

effects). The negative coefficient on 'Treated*PolicyOn' implies an improvement in the treatment group's depression scores relative to the control group, but this is not significantly different from zero. The match between the treatment and control groups (in terms of the depression score) is shown by the coefficient on the 'Treated' variable. As suggested by the descriptive statistics in Table 4, pupils in the combined treatment group score on average 0.08 of a standard deviation worse than those in the control group on the depression inventory (given this specification). The two asterisks following the coefficient indicate that it is significant at 5%, i.e. there is only a 5% chance that it could have arisen by chance if in fact the treatment and control values were equal. [19]

This coefficient on 'Treated' increases slightly when controls for pupils' demographic characteristics are added, but drops and becomes insignificant once school fixed effects are included. This is because, as we suggested above, many of the differences between pupils in the treatment and control groups here are due to between-school differences rather than within-school differences, so once we control for a pupil's school the difference disappears.

The coefficient on 'Treated*PolicyOn' gets slightly larger as more controls are added, implying that there is an average improvement as a result of the workshops once we account for the differences between pupils in the control and treatment groups. Columns 4 and 8 present specifications using pupil fixed effects. This works by using multiple observations of the same pupil over time to control for all the characteristics of the pupil which are fixed through time. This should control for many of the differences in composition between the treatment and control groups, such as gender, prior attainment etc. For this reason it is our preferred specification. Based on this, we obtain a programme impact of 0.07 of a standard deviation in the short run, significant at 10% (Column 4), with no detected effect at one-year follow-up (Column 8). This represents a small programme effect.[20]

We know that there are differences in composition between the different treatment groups, so it is worth examining treatment effects separately by treatment group using this econometric specification, which allows us to control for pupil characteristics. Table 10 (compare Table 30 of Challen at al, 2009) presents difference-in-differences estimates of the treatment effects with the three groups of workshops entered separately. Again, the coefficient on 'Treated*PolicyOn' gives the difference-in-differences estimate, i.e. the average impact of being in the treatment group on a pupil's (standardised) depression symptoms score at the end of the treatment. The first four columns present different specifications of the post-treatment score (pupils' scores in 2008); while columns 5-8 contain specifications where the outcome variable is the depression symptoms score measured at one year follow-up (June 2009).

As suggested by the simple comparison of baseline characteristics, the start-mid year treatment group (Group 1, top panel) seems well-matched to the control group in terms of the depression symptoms score, hence the absence of a significant coefficient on the

[19] Here and in the tables below, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

[20] In specifications using pupil fixed effects we also control for mean reversion, by including variables indicating the initial level of symptoms of the outcome variable, and for school trends.

'Treated' variable in the first panel in both time periods. Moreover, adding controls to the regression does not change much the size of the coefficient on 'Treated*PolicyOn' for this group of workshops, implying that the workshop and control groups are reasonably well balanced on the control characteristics too, although the estimated treatment effect does become slightly larger (more negative) and more significant as controls are added to the specification. Based on this group of workshops, we obtain a treatment effect of 15% of a standard deviation improvement in pupils' depression scores in the short run when pupil fixed effects are added in column (4), a small-to-moderate treatment effect. [21] [22] This has disappeared by the one-year follow-up, where the estimated effect is basically zero.

The second panel gives the same information for the second group of workshops. The problem of the mismatch between the treatment and control groups for this experiment is shown by the positive and significant coefficient on the 'Treated' variable: this means that pupils who were in the treatment group scored higher (worse) on the depression symptoms inventory even before the treatment had started. This implies that the as-if randomisation did not successfully create comparable treatment and control groups, and so taking comparisons of these groups before and after the workshops will not provide a true estimate of the treatment effect. The importance of controlling for pupil characteristics is suggested by the change in the coefficient on 'Treated*PolicyOn' for the Group 2 workshop sample, when pupil fixed effects are added in column 4: this reduces the estimate of the programme impact to zero.

Both of the first two treatment groups show an improvement in depression scores, although this only remains significant for Group 1 workshops in the short run (at 1%), once pupil fixed effects have been included. Interestingly, the measured impact of the Group 2 workshops appears to have increased by one-year follow-up, with a point estimate of 0.2 of a standard deviation when pupil fixed effects are included (column 8, panel 2). It is possible that the impact of an intervention could increase over time. Reasons for this include better use of the programme skills over time as pupils gain more practice; a 'sleeper' effect where an impact on well-being is only seen with a lag; or the programme preventing a worsening in the treatment group which does appear in the control group. While these may all be plausible explanations for the difference, given the mismatch between treatment and control at baseline we are suspicious of the estimated impact for the Group 2 workshops, as this effect might partly be the result of mean reversion (although we are already controlling for this in the specifications containing pupil fixed effects).

These effects compare with a positive coefficient on 'Treated*PolicyOn' (i.e., a worsening of scores in the treatment group relative to the control group) in all specifications for the short run for the start-end year treatment group (Group 3). The short-run programme impact for this set of workshops is zero, as only the coefficient in the first (uncontrolled) specification is significant. However, at one-year follow-up this has grown to about 0.1 of a standard deviation, and is significant at 10% even when pupil fixed effects are included. This implies that pupils who were in this treated group had actually got worse relative to the control group by one year after the end of the workshops. Again, we suspect that this result might be partly

[21] Pupil fixed effects control for characteristics of pupils that are fixed over time.
[22] Remember that a higher score indicates more and worse symptoms, so a negative coefficient implies an improvement in these regressions.

due to mismatch between the treatment and control groups: the coefficient on 'treated' for these workshops jumps around as the specification changes, implying that they are not matched on a number of characteristics, and that these characteristics have some relationship with the outcome variable. Indeed, Table 7 suggests that the treatment and control groups differ in terms of demographic characteristics.[23] This is in contrast to the Group 1 workshops, for which the coefficient on 'Treated' is always small and insignificant and does not change much with different specifications.

Nevertheless, we see clear treatment heterogeneity by the three groups of workshops: the p-value of a chi-squared test of constant treatment effect is less than 0.1 for all specifications, meaning that it is unlikely that such different estimates of the treatment effect would arise by chance. Combined with the evidence from Tables 4-7, this seems to suggest that although the 'as-if' randomisation may have worked for the Group 1 workshops, it was not successful in the other two samples. We know that some schools overtly or covertly decided to select pupils to participate in workshops, rather than randomly assigning them as had been agreed, and that this was particularly evident in the mid-end year treatment group.[24] The positive treatment effect for the start-end year workshops at one-year follow-up is harder to explain, although a treatment effect of zero combined with differential trends between schools, including adverse attrition of the evaluation sample, could be responsible.

We suggest that the as-if randomisation worked well enough for the first out of these three workshop sets, but apparently failed on the other two to some extent. We therefore conduct the remainder of the quantitative analysis using both the full sample, and the Group 1 sample plus the control group, as we believe that this provides a more reliable picture of the impact of the workshops: if the results are robust to using these two samples, we can be more confident that our estimates are correct.

This is not to say that there might not be differential effects caused by different frequency and timing: fortnightly workshops (as in Group 3) were extremely unpopular with teachers, as it proved difficult to teach with such a long gap between lessons.[25] In comparison, Group 1 and 2 workshops were held weekly with no obvious problems relating to timing. For this reason it might be sensible to consider workshops from Groups 1 and 2 as one type of programme, while Group 3 workshops constituted a programme that was slightly different. Moreover, it is possible that since pupils are new to the school in September 2007 when the weekly Group 1 workshops begin, these might have had a greater potential for impact than the weekly Group 2 workshops which began in February because pupils could be easier to influence. School transitions can be difficult and can have an adverse impact on subjective well-being; starting the programme when pupils first came into the school might help to mitigate these problems more effectively than waiting until half way through the academic year. If this is the case, we would expect a larger programme impact for workshops starting

[23] Although note that once school fixed effects are included in columns 3 and 7 the coefficient on 'treated' falls to zero for this group, supporting the hypothesis that the differences between treatment and control pupils are in large part due to their originating from different schools.
[24] Since pupils were new to the school in September 2007, even if schools did attempt to select pupils for workshops on the basis of psychological distress they would have been less successful at this point as they did not know the pupils well enough. By February 2008 they might have known pupils well enough to select more effectively.
[25] See Challen et al. (2009), p86.

earlier in the year, such as the Group 1 and Group 3 workshops. These two mechanisms combined could result in weekly workshops beginning in September as being the most effective, and that appears to be what we find. However, we have no way of disentangling workshop timing and frequency from a host of other variables that could impact the treatment effect, so we can only speculate as to why there appears to be treatment heterogeneity between the three groups of workshops.

Heterogeneity of impact by pupil characteristics

Now we can go on to examine heterogeneity in the impact of treatment by pupil characteristics, essentially, whether different groups of pupils appear to be affected more or less by the workshops. We run the same specifications on our two different samples to check for the robustness of the results.

Table 11 presents evidence on the heterogeneity of treatment effects by pupil characteristics, using the full sample. These results are generated from separate specifications relative to controls, each column within a box representing a separate regression. Here we use the full sample, which is larger than the sample used in the 2009 evaluation report, and includes all three treatment groups. The specifications we use are slightly different, but columns 1-3 of Table 11 are roughly comparable with Table 31 of the 2009 report (although Table 31 of the 2009 report only uses the Group 1 treatment, so is more similar to Table 12 of the current report). These columns present the combined estimated treatment effect immediately after the end of the start-mid year workshops (around February-April 2008) and at the end of the academic year (July 2008). Columns 4-6 of Table 11 update this by presenting the estimated impact in June 2009.

We have changed the specification we are using because mean reversion is a serious challenge in estimating programme impact, especially with subjective well-being measures. Specifically, high scores recorded at a point in time are very likely to fall through time, and this could be interpreted as an impact of the programme unless this is controlled for in both the treatment and control groups. For this reason we have included controls for the initial level of symptoms, and so estimated reductions in depressive symptoms score attributed to treatment should not be caused simply by mean reversion. We also include 22 different school trends (one per school) in order to account for differential change through time by school.

These results compare to the overall average treatment effects in Table 9 of -0.068 (0.036) for the short-run impact and -0.007 (0.044) for the one-year follow-up.

For this sample we cannot reject that the size of the treatment effect on girls and boys is equal (p=0.63).[26] However, pupils with special educational needs seem to benefit significantly more from the workshops than those without SEN, and the difference between the two groups is stronger for boys. Moreover, the treatment effect appears to persist to the

[26] The p-values reported are from a test of equality of the coefficients immediately above: if the p-value is greater than 0.1 this implies that the coefficients are not significantly different from each other and therefore that pupils with (e.g.) FSM entitlement are not more or less likely to benefit from the workshops than those without.

one-year follow-up for boys with SEN (for girls the coefficient is equally large at this point, but only significant at 15%). There appears to be no difference by entitlement to free schools meals. When we split the sample by Key Stage 2 attainment, to compare the treatment effect for pupils who did or did not attain the national target of Level 4 in English or maths, we find that girls who did not attain the national targets show a much stronger treatment effect than girls who did. Moreover, this persists until the one-year follow-up. There is no difference for boys by prior attainment. Splitting the sample by the baseline quintile of the depression score, it appears that much of the average effect comes from an improvement in those who were in the worst scoring quintile at baseline (5^{th} quintile), notably girls.

As we showed above, the full sample is not balanced by outcome or demographic variables at baseline (see Table 4). Since there appear to be differential treatment effects by pupil characteristics, any differences in the composition of the treatment and control groups could result in an unrepresentative estimate for the average treatment effect. For this reason we also present these same regressions on treatment heterogeneity for the Group 1 workshops in Table 12 for which treatment and control groups are well balanced at baseline.

Table 12 is similar to Table 31 of the 2009 evaluation report, as it presents the programme impact using only the Group 1 workshops and the control group, adding results for the one-year follow-up.[27] These results compare to the overall average treatment effects in Table 10 of -0.141 (0.049) for the short-run impact and -0.067 (0.061) for the one-year follow-up.

Because of the change in sample size and specification, some of the differences in programme impact reported last year are no longer evident. There now appears to be an average treatment effect for both girls and boys during the first year (though not at the one-year follow-up), and although this is larger and more significant for girls the p-value from the test of equality of the two coefficients (p=0.27) suggests that they are not significantly different from each other.

Further disaggregating by pupil characteristics, we find that pupils with some form of SEN are significantly more likely to improve their depression scores relative to the control group as a result of being in workshops than pupils without SEN, although there is no significant difference in treatment effect between girls with or without SEN. There is apparently no difference in impact by free school meals eligibility (FSM).

However, it is interesting that differences between some subgroups appear to be larger at the one-year follow-up than immediately post workshops. It appears that girls who did not achieve level 4 English (or maths) at Key Stage 2 (KS2) and who participated in workshops improved their depression scores by about half of a standard deviation compared to the control group (a moderate-to-large programme effect), and that (unusually) this effect persists at one-year follow-up. This is significantly different from treated girls who did attain the national targets at KS2, who obtained an improvement of 0.2 of a standard deviation in depression scores relative to the control group, which had faded out by one-year follow-up. There was no such difference for boys.

[27] However, as discussed above, Table 12 does use a larger sample than Table 31, and the specifications we use this year are slightly different from the ones in Challen et al. (2009).

Once we split the sample by the quintile of the baseline (September 2007) depression score, we find important differences in measured impact by baseline score. Pupils who scored in the worst (highest) 20% of scores in September 2007 (quintile 5) improved significantly relative to the control group, and the effect was particularly large (70% of a standard deviation) for girls in the 5th quintile. This improvement is partly maintained at one-year follow-up for girls, although for boys in the fifth quintile there is now zero effect. What is striking is that there is no overall effect across the other quintiles. This could partly be because of small sample sizes meaning that there is low power to detect small effects, but it also appears likely that the measured gains from the programme are concentrated among those who had the worst subjective well-being at the start of the programme. Note that this is unlikely to be due to mean reversion, as we have controlled for this in the specification.[28] The Penn Resiliency Program was originally conceived as a depression prevention programme for targeted (at-risk) adolescents, so it is perhaps not surprising that pupils closer to this category might respond the most. However, given the strong ceiling effect imposed on pupils who obtain low (good) scores at baseline due to the skewed distribution of the depression scale we are using, it is also possible that pupils with higher scores simply have further to go in terms of possible improvement. That is to say, the scale we are using to measure symptoms of depression is only sensitive around the part of the distribution where there are some symptoms of depression, and not sensitive to improvements in children who are already reporting few or no symptoms.

Taking the results from these two samples together, we suggest that there is likely to be some treatment heterogeneity by SEN status, prior attainment, and baseline depression score, with lower attaining and initially more depressed children benefitting more. Given that the largest treatment effect is found when the sample is split by baseline depression score it is possible that the large effects found for lower attaining pupils could be due to the fact that they are more likely to have poor psychological scores, rather than any direct effect of SEN or low attainment. We test this hypothesis in the following way: we further split the treated sample into four groups (pupils with low initial depression scores and low maths attainment; pupils with low depression and high maths; pupils with high initial depression and low maths; and with high depression and high maths), and run the same regressions. We find that there is a significant treatment effect for each of these groups except pupils with low depression and high maths (regressions not shown; this result applies whether using the full sample or the Group 1 treatment only). This implies that controlling for baseline depression score there is a differential impact on depression scores for low-attaining pupils over and above that which might be due to their starting with a worst depression score.

Overall, the analysis of the depression scores suggests treatment heterogeneity by organisation/timing of treatment, part of which may have been caused by selection of pupils into the workshop groups, although it is also possible that the timing and frequency of the workshops could have an impact. There is also significant heterogeneity of treatment impact by pupil characteristics, with lower attaining pupils and those who began the year with worse depression scores apparently gaining more from workshops in the short run. At one-year

[28] One result of including controls for mean reversion was that the counterintuitive positive treatment effect for 1st quintile boys that we found in the first report disappears (see Challen et al. 2009, Table 31). As we suspected, this result was due to pupils scoring 0 or 1 at baseline being unable to improve relative to this.

follow-up the only impacts to persist are those for low-attaining girls and those with the worst baseline for subjective well-being.

Programme impact on symptoms of anxiety

Table 13 is the equivalent of Table 8, but uses the anxiety (Revised Children's Manifest Anxiety Scale Score - RCMAS) score as the outcome variable. The general pattern of results is similar to that seen in Table 8 for the depression symptoms score, which is not surprising given that these two scores are strongly correlated. However, Table 13 (raw scores by treatment group) and Table 14 (pooled econometric specification) suggest that there is no impact of treatment on anxiety score on average. [29] In Table 14, the significant and positive coefficient on 'Treated' suggests, as in Table 4, that the treatment group has a worse anxiety score at baseline than the control group, implying that they are not well matched along this dimension. Table 15 unpacks this further for the three treatment groups separately.[30] As for the depression score, the difference between the treatment and control groups for the start-mid year treatment is small and insignificant, while those treated from the middle of the year to the end score significantly higher than the control group at the baseline and the mid-year measurement dates. Only for the start-mid year treatment group is there a significant treatment effect, and this has disappeared by the one-year follow-up. As for the depression score, we see similarly peculiar results for the Group 2 and Group 3 treatment groups at one-year follow-up, with the Group 2 treatment showing a strong improvement and the Group 3 workshops showing a significant deterioration in anxiety scores relative to the control group. However, unlike the pattern for the depression score, once the pupil fixed effects model is used in column 8 of Table 15, the estimate of programme effect for all three treatments is essentially zero. This suggests that it is pupil characteristics that are driving the estimated treatment effects, rather than participation in workshops.

Table 16 shows the equivalent specifications with the anxiety score as an outcome as are shown in Table 11 for the depression score, using the full sample with all three treatment groups. Interestingly, here it is boys who show a greater reduction in scores, particularly boys with SEN or an entitlement to FSM. A similar effect on girls with low prior attainment that we saw for depression appears here, despite the overall zero effect on anxiety symptoms for girls. Children who scored in the worst quintile of the anxiety score at baseline appear to show a greater reduction as a result of the workshops, but this is not strong enough to appear at a reasonable level of significance. None of these effects lasts beyond the short run.

Table 17 presents the same specifications using the Group 1 treatment sample.[31] The results are broadly similar here, although the impact is generally stronger, and the effect even persists to one-year follow-up for the groups that appeared to benefit more in the short run: boys with FSM or SEN, and lower attaining girls. When the sample is split by the quintile of the baseline anxiety score we only observe effects in the bottom two quintiles.

[29] The equivalent in the 2009 evaluation report is Table 32.
[30] The equivalent in the 2009 report is Table 33.
[31] The equivalent in the 2009 report is Table 34.

Thus the results seem broadly consistent between the two samples: overall, any impact on anxiety is more muted than the impact on depression, and is concentrated in a few groups of pupils: boys, particularly boys with SEN or FSM entitlement, and lower-attaining girls.

Absence from school

Annual absence from school is measured as the fraction of school sessions for which pupils were absent during the academic years 2006-7 and 2007-8. This is the sum total of authorised and unauthorised absences during the year. Since workshop pupils participated in workshops in the academic year 2007-8, if there is an effect of the workshops on attendance we would expect to see fewer absences for these pupils relative to the control group. The data on absence was obtained from the National Pupil Database, and is likely to be less subject to reporting biases than the psychological or questionnaire measures we have used in the analyses above. In addition, since we can obtain data on a child's attendance at school for as long as they attend a state school in England, there is a much lower level of attrition than for the other measures. There is also much less danger of selective attrition from not filling in the questionnaires due to absence or refusals, and so we can be less concerned about the possibility of sample selection bias. However, one disadvantage of using these data is that data for the academic year 2008-9 was not available at the time of writing and so we are only able to evaluate the short-run impact to July 2008.

Table 18 is the equivalent of Tables 9 and 14, presenting estimated treatment effects for the pooled workshop groups with absence as the outcome variable. Here the pooled sample appears balanced between treatment and control groups on this outcome, with the coefficient on 'Treated' small and insignificant. The coefficient on 'Treated*PolicyOn' is not significant until the specification using pupil fixed effects is used in Column 4, but here we obtain an estimate of programme impact on attainment that is 18% of a standard deviation and significant at 1%.

Table 19 presents the same specifications with the treatment split into the three groups. Interestingly, none of the treatment and control groups seems well-balanced here, with treated pupils showing significantly higher baseline levels of absence than control pupils for the first two workshop groups, and significantly lower ones for the third. However, this is probably largely due to pupils originating from different schools: once school fixed effects are included in column 3 the coefficient on 'treated' is greatly reduced for all three groups and is no longer significant. If the programme reduces absence we should see a negative coefficient on the Treated*PolicyOn variable, as indeed we do for all three groups, although this is only significant for Group 1 workshops. However, using the column 4 pupil fixed effects specification we cannot reject the hypothesis that all three treatment effects are equal ($p=0.30$).[32]

[32] The full sample is well-matched on the outcome variable between treatment and control groups here, but we continue to include results for both the full sample and the Group 1 workshops in order to enable comparison with the other outcomes, and because the full sample is still not balanced on other pupil characteristics which could affect the results.

Examining the treatment effect by pupil characteristics for the full sample shows no overall pattern of differential treatment effect by the characteristics we use to split the sample (Table 20). However, using only the Group 1 workshops (Table 21) there does seem to be a larger impact on pupils entitled to free school meals and pupils who did not attain KS2 maths at level 4, particularly for boys in these categories, for whom the treatment effect is at least 0.5 of a standard deviation. Splitting the sample by quintiles of baseline absence suggests that the impact does not depend on the initial level of absence, but is fairly constant across the distribution. So overall, we find an average treatment effect of 0.18-0.25 of a standard deviation improvement in absence rates, and find that this effect is fairly homogeneous across different groups of pupils.

To think about what kind of impact this is in practice, we can use the data from the evaluation sample. The mean absence rate for these pupils in 2006-7, when they were in the final year of primary school, was about 0.04 or a 4% absence rate. This had risen to 0.07 in 2007-8. The standard deviation in which the effect is measured is that of the absence rate in 2006-7, which has a standard deviation of about 0.06. The median number of possible school sessions for these pupils in 2006-7 (for which attendance is recorded) was 308 (306 in 2007-8). If UKRP workshops reduced absence rates by .2 of a standard deviation on average, this is equivalent to 0.06*.2*308=3.7 more sessions attended over the course of the year, or about 1.8 more school days. This might not seem like a large improvement with respect to the total number of sessions possible (it is only 1.2% of the 308 possible sessions for the year), but it is sizeable relative to the overall absence rate. Relative to the 7% absence rate in 2007-8 it represents an improvement of 0.17 or 17%.

It is worth noting that whereas the psychological measures are assessed at a point in time, and can therefore be deemed to be ‘pre’ and ‘post’ measures for the workshops, the absence measure is accumulated over the course of the year and so includes the time period when pupils in the treatment group were attending workshops. Pupils generally enjoyed the workshops, and some teachers commented that attendance was higher on UKRP days because pupils did not want to miss the lessons. If this is the case, it is possible that the effect observed could be entirely due to increased attendance on days with UKRP lessons. This is important both because one would expect the effect to wear off as soon as the workshops finished, and moreover we might not see so much value in increased attendance at school if pupils only attend this additional ‘fun’ lesson rather than other subjects as well. We will soon have data from the academic year 2008-9, and so will be able to see whether the effect persists. In response to the second concern, a session will correspond to at least one lesson, usually two or three, so even if pupils were only motivated to attend school for one lesson it is likely that they attended others as well. It is also possible that attending school more often might be good in itself, promoting connection to the school and ensuring that pupils are not spending time outside of school where they may be vulnerable.

Conclusion

In this follow-up quantitative analysis we use both the full evaluation sample and the experiment that appears to have well-balanced treatment and control groups to examine programme impact.

There appears to be an overall short run impact on depression scores and attendance, although this has faded by one-year follow-up for the depression score (we have not yet been able to examine this for attendance due to lack of data). There was also a short-run impact on anxiety, but this was smaller, and concentrated in a few groups of pupils: boys, particularly boys with SEN or FSM entitlement, and lower-attaining girls. We find some impact of the programme on depression and anxiety symptom scores at one-year follow-up, although there has been a general fading out of the effects of the programme by this point. We find significant heterogeneity in the measured effects of the UK Resilience Programme, in terms of how the workshops were organised, the outcomes assessed (depression, anxiety and absence), and the impact on different groups of pupils. Workshops that were timetabled weekly and which started at the beginning of the academic year had a significantly larger impact on depression and anxiety scores than those which either started later or were timetabled fortnightly. (They also had a larger impact on absence rates, though this was not significantly different from the other workshop groups.)[33]

Furthermore, we found variation in the impact of treatment by pupil characteristics: in general, lower attaining and more disadvantaged pupils appear to gain more from the workshops, and in some cases the programme impact has not faded by one-year follow-up for these groups. Specifically, the impact of the workshops was larger for pupils with Special Educational Needs when the outcome was anxiety or depression; for pupils entitled to free school meals (anxiety and attendance); for pupils who had not attained the national target levels in Key Stage 2 exams (depression, anxiety and attendance); and for pupils with worse initial scores for symptoms of depression or anxiety (depression and anxiety). Interestingly, there did not seem to be so much heterogeneity in the impact of workshops when the outcome was the absence rate.

Some of the observed heterogeneity might be at least partly due to the lack of sensitivity of the psychological measures we are using: they are good at discriminating between children above a certain level of symptoms, but are unable to detect improvements in those who already have good psychological well-being. This is one reason why the result we obtain when the absence rate is the outcome is particularly interesting: most children have at least some days off school, so even though this measure also has a very skewed distribution there is still perhaps more room for detecting an effect across the distribution. However, we will have to wait until we obtain later years of data to see whether this impact lasts beyond the programme year.

Further study of the longer run effects of the workshops will be possible after the next wave of data has been collected in June 2010.

[33] We should bear in mind that the association of these workshops with a greater programme impact is not necessarily causal.

Implementation and policy issues

Here we present some possible policy implications based on the results reported above and from the findings of last year's report. Many of these points are not recommendations, but they do highlight issues to consider when implementing the programme, based on the evaluation results and the results of prior PRP trials. They are to a certain extent speculative, as we cannot know how the workshops would have functioned had they been organised differently. However, we provide a discussion of what we believe could be points for consideration if intending to use the UKRP in schools or elsewhere.

Workshop timing

Teaching fortnightly workshops was very unpopular with teachers (Challen et al., 2009, p87). Moreover, the preliminary quantitative work suggests that these workshops had a smaller impact than those that took place once a week, although we cannot be sure that this was due to timing and not other factors. Fortnightly workshops were also less popular with pupils, though again it is not clear that this was causal (Challen et al., 2009, p21). All previous PRP trials have scheduled workshops to take place at least once a week. Timetabling workshops once a week rather than once a fortnight therefore seems preferable, if this is possible.

Class sizes and active ingredient

About 20% of pupils in workshops in the first year of UKRP were in groups of 16 or more, though only 3% were in groups of 18 or more. Facilitators thought that small class sizes were very important to the success of the programme (Challen et al., 2009, p37). It might therefore be important to maintain small classes in order to maintain the quality of the programme. Previous Penn research finds that programme quality is an important factor in predicting programme impact, and many previous trials have used classes smaller than 15.

Logistical issues, staff support, and child protection

One of the most persistent issues arising in the qualitative work was the importance of internal school organisation. The demands of the programme in terms of staffing and rooms are clear (halving the size of a class necessitates twice as many rooms and twice as many teachers), but in addition the need for support for facilitators not accustomed to dealing with issues around safeguarding children came up. Despite the insistence in the curriculum that pupils do not have to discuss their problems in class (rather, discussion should focus on how one could solve example problems), the subject matter does often lead to disclosures and a number of pupils choose to confide in their group facilitator during or after the lessons. As well as dealing with inappropriate disclosures within the classroom, facilitators have to support the burden of expectation from pupils who want solutions, and quite simply the sadness of some of the stories they are told. It might therefore be advisable to ensure before the start of the programme that facilitators know how child protection arrangements work

within their school so that both they and pupils can obtain appropriate support.[34] Formal support similar to that provided for social care professionals has also been suggested, but a more feasible way of providing support might simply be to have more UKRP facilitators within each school.

It is also worth considering the impact that running the workshops could have on the number of child protection cases coming to light within schools. The surveys that schools have been administering for the evaluation have revealed children reporting serious psychological symptoms who would not normally have been noticed (we feed back to schools on any child protection issues that arise from the surveys). Making sure school staff are prepared for this could be an important part of preparation for the workshops.

Heterogeneous effects and pupil selection

Certain groups of pupils appeared to benefit more from the workshops, particularly those who did not achieve the national targets at Key Stage 2, and those who started the school year with high levels of depression or anxiety symptoms. This might seem to suggest that these pupils should be given priority to be included in workshops over pupils who are achieving national targets or who do not display obvious signs of anxiety or depression. However, it is worth bearing in mind the following points:

- The measured impact on these pupils is the impact of the programme delivered to 'universal' or mixed workshop groups, not of groups consisting entirely of targeted pupils. One cannot therefore assume that the same impact would be obtained if workshop groups were targeted.

- Some schools that did run workshops entirely with targeted pupils reported these as being very difficult to manage and not very successful, compared to those that had pupils arbitrarily allocated to the group or a mixture of arbitrarily allocated and targeted pupils (see Challen et al., 2009, p35).

- The same applies to levels of academic attainment: many facilitators commented that SEN groups or lower set groups did not go well, or that the presence of more able or more literate pupils aided the success of the lessons (Challen et al., 2009, p35).

- It is interesting that facilitators and other school staff often appeared to assume that higher ability pupils were naturally more resilient, or had fewer problems. Yet almost all facilitators claimed to use the UKRP skills themselves, so it seems unlikely that higher ability pupils would be unable to benefit from the skills if facilitators claimed to.

- Even if pupils were to be targeted for inclusion in workshops, it is important that they should be targeted appropriately. Previous research suggests that school staff tend to identify pupils with behaviour problems rather than those with emotional difficulties, yet the programme is primarily designed to address the latter. The process of

[34] We should stress that we did not come across any staff in the evaluation schools who were not aware of their schools' child protection arrangements.

targeting would need to be carefully considered, or it could be in danger of becoming intrusive and inaccurate.

- Participation in programmes perceived to be targeted and remedial can attract stigma for those who participate. Universal programmes avoid this type of labelling.
- The measures used in the evaluation are sensitive to differences in the severity of symptoms of depression and anxiety, but are not good at distinguishing between children who have little or no symptoms. For instance, they would not be able to detect any improvements in well-being for pupils who showed no initial symptoms of depression, although this would not necessarily mean that these children did not benefit.
- The skills pupils used most (as reported by both pupils and facilitators) were the interpersonal skills around negotiation and assertiveness, and techniques for self-control. Since all pupils are likely to experience conflict and problems around everyday social interactions it is likely that all pupils could benefit from the workshops, at least in these areas.
- It is worth bearing in mind that the treatment effects we estimate are averages; it is possible that some pupils will get nothing from the programme and that some might even find it distressing.
- The analysis conducted above is preliminary: we have not yet examined all outcome measures and so do not have a full picture of the impact of the programme.

Given these considerations, it is not clear that the most effective way of delivering workshops would be to select pupils perceived to be most at need, given that these pupils may not in fact be those who would benefit most (through poor selection); that workshops might not work well even if pupils were well selected, as targeted groups have tended to be less successful; and that pupils might attract stigma from being in the workshops. Facilitators reported that more able pupils accessed the materials more easily, and that workshop groups went better with them (Challen et al., 2009). However, given that the quantitative work finds a smaller impact on more able or higher attaining pupils it is also not clear that these pupils should be given priority either.

If classes were targeted it would be worth thinking about how to do this sensitively, how to avoid stigma (e.g. withdrawing children from classes as for other extracurricular activities), and how to put together pupils who could function well as a group.

Table 4: Comparison of treatment and control group baseline means – Full sample

Full sample: all experiments pooled	Treatment group	Control group	p-value of test of equality of means
All treatments			
Depression score at baseline	**8.55**	**8.03**	0.0269
Standard deviation	6.97	6.14	
number of observations	1607	1546	
Anxiety score at baseline	**9.38**	**8.97**	0.0838
Standard deviation	6.81	6.30	
number of observations	1585	1525	
Pupil-reported behaviour score at baseline	**11.00**	**10.81**	0.4012
Standard deviation	6.32	6.20	
number of observations	1586	1522	
Teacher-reported behaviour score at baseline	**5.37**	**5.52**	0.4864
Standard deviation	5.45	5.42	
number of observations	1429	1197	
Life satisfaction score at baseline	**33.71**	**34.19**	0.0320
Standard deviation	6.19	5.99	
number of observations	1539	1454	
Fraction of sessions absent (2006-7)	**0.04**	**0.04**	0.7408
Standard deviation	0.06	0.06	
number of observations	1901	1909	
Gender (male=0; female=1)	**0.49**	**0.47**	0.1810
Standard deviation	0.50	0.50	
number of observations	1607	1546	
Special Educational Needs	**0.24**	**0.27**	0.0317
Standard deviation	0.43	0.44	
number of observations	1607	1546	
Free School Meals	**0.22**	**0.30**	0.0000
Standard deviation	0.41	0.46	
number of observations	1607	1546	
KS2 English score	**4.07**	**3.91**	0.0000
Standard deviation	0.85	0.93	
number of observations	1561	1502	
KS2 maths score	**4.05**	**3.90**	0.0000
Standard deviation	0.86	0.91	
number of observations	1561	1502	

Notes: this table presents baseline means for a range of variables to enable us to gauge the similarity of the treatment and control groups. Here we present means for the full sample: all three treatment groups pooled plus the control group. The last column gives the p-value of a test of equality of means between the two groups: if this is less than 0.1 for a variable the treatment and control groups are said to be significantly different from one another at the 10% level. This means that there is less than a 10% chance that an equally large difference between the two would have arisen by chance if they were in fact the same, i.e. that the two groups are not well matched on this variable at baseline.

Table 5: Comparison of treatment and control group baseline means by workshop timing – Group 1 workshops (weekly start-mid year workshops)

Group 1 workshops plus control group	Treatment group	Control group	p-value of test of equality of means
Treatment: start year - mid year			
Depression score at baseline	**8.27**	**8.03**	0.5217
Standard deviation	7.16	6.14	
number of observations	373	1546	
Anxiety score at baseline	**9.37**	**8.97**	0.2811
Standard deviation	6.86	6.30	
number of observations	365	1525	
Pupil-reported behaviour score at baseline	**10.70**	**10.81**	0.7503
Standard deviation	6.13	6.20	
number of observations	367	1522	
Teacher-reported behaviour score at baseline	**5.21**	**5.52**	0.4050
Standard deviation	6.12	5.42	
number of observations	289	1197	
Life satisfaction score at baseline	**33.80**	**34.19**	0.2747
Standard deviation	6.24	5.99	
number of observations	351	1454	
Fraction of sessions absent (2006-7)	**0.05**	**0.04**	0.0004
Standard deviation	0.07	0.06	
number of observations	446	1909	
Gender (male=0; female=1)	**0.47**	**0.47**	0.9735
Standard deviation	0.50	0.50	
number of observations	373	1546	
Special Educational Needs	**0.23**	**0.27**	0.1113
Standard deviation	0.42	0.44	
number of observations	373	1546	
Free School Meals	**0.27**	**0.30**	0.2445
Standard deviation	0.44	0.46	
number of observations	373	1546	
KS2 English score	**3.99**	**3.91**	0.1471
Standard deviation	0.95	0.93	
number of observations	366	1502	
KS2 maths score	**4.03**	**3.90**	0.0139
Standard deviation	0.93	0.91	
number of observations	364	1502	

Notes: this table presents baseline means for a range of variables to enable us to gauge the similarity of the treatment and control groups. Here we present means for the Group 1 workshops plus the control group. Group 1 workshops were those held weekly that started at the beginning of the academic year. The last column gives the p-value of a test of equality of means between the two groups: if this is less than 0.1 for a variable the treatment and control groups are said to be significantly different from one another at the 10% level. This means that there is less than a 10% chance that an equally large difference between the two would have arisen by chance if they were in fact the same, i.e. that the two groups are not well matched on this variable at baseline.

Table 6: Comparison of treatment and control group baseline means by workshop timing – Group 2 workshops (weekly mid-end year workshops)

Group 2 workshops plus control group	Treatment group	Control group	p-value of test of equality of means
Treatment: mid year - end year			
Depression score at baseline	**10.64**	**8.03**	0.0000
Standard deviation	8.14	6.14	
number of observations	175	1546	
Anxiety score at baseline	**11.47**	**8.97**	0.0000
Standard deviation	7.34	6.30	
number of observations	172	1525	
Pupil-reported behaviour score at baseline	**12.64**	**10.81**	0.0004
Standard deviation	6.92	6.20	
number of observations	166	1522	
Teacher-reported behaviour score at baseline	**6.13**	**5.52**	0.2451
Standard deviation	5.64	5.42	
number of observations	116	1197	
Life satisfaction score at baseline	**33.10**	**34.19**	0.0323
Standard deviation	6.91	5.99	
number of observations	159	1454	
Fraction of sessions absent (2006-7)	**0.06**	**0.04**	0.0000
Standard deviation	0.07	0.06	
number of observations	229	1909	
Gender (male=0; female=1)	**0.43**	**0.47**	0.2878
Standard deviation	0.50	0.50	
number of observations	175	1546	
Special Educational Needs	**0.28**	**0.27**	0.8004
Standard deviation	0.45	0.44	
number of observations	175	1546	
Free School Meals	**0.31**	**0.30**	0.8454
Standard deviation	0.46	0.46	
number of observations	175	1546	
KS2 English score	**3.98**	**3.91**	0.3414
Standard deviation	0.93	0.93	
number of observations	168	1502	
KS2 maths score	**4.00**	**3.90**	0.1723
Standard deviation	0.92	0.91	
number of observations	168	1502	

Notes: this table presents baseline means for a range of variables to enable us to gauge the similarity of the treatment and control groups. Here we present means for the Group 2 workshops plus the control group. Group 2 workshops were those held weekly that started in the middle of the academic year. The last column gives the p-value of a test of equality of means between the two groups: if this is less than 0.1 for a variable the treatment and control groups are said to be significantly different from one another at the 10% level. This means that there is less than a 10% chance that an equally large difference between the two would have arisen by chance if they were in fact the same, i.e. that the two groups are not well matched on this variable at baseline.

Table 7: Comparison of treatment and control group baseline means by workshop timing – Group 3 workshops (fortnightly start-end year workshops)

Group 3 workshops plus control group	Treatment group	Control group	p-value of test of equality of means
Treatment: start year - end year			
Depression score at baseline	**8.31**	**8.03**	0.2790
Standard deviation	6.63	6.14	
number of observations	1059	1546	
Anxiety score at baseline	**9.04**	**8.97**	0.8008
Standard deviation	6.65	6.30	
number of observations	1048	1525	
Pupil-reported behaviour score at baseline	**10.85**	**10.81**	0.8853
Standard deviation	6.26	6.20	
number of observations	1053	1522	
Teacher-reported behaviour score at baseline	**5.32**	**5.52**	0.4002
Standard deviation	5.22	5.42	
number of observations	1024	1197	
Life satisfaction score at baseline	**33.78**	**34.19**	0.0920
Standard deviation	6.05	5.99	
number of observations	1029	1454	
Fraction of sessions absent (2006-7)	**0.03**	**0.04**	0.0000
Standard deviation	0.05	0.06	
number of observations	1226	1909	
Gender (male=0; female=1)	**0.51**	**0.47**	0.0318
Standard deviation	0.50	0.50	
number of observations	1059	1546	
Special Educational Needs	**0.23**	**0.27**	0.0299
Standard deviation	0.42	0.44	
number of observations	1059	1546	
Free School Meals	**0.18**	**0.30**	0.0000
Standard deviation	0.39	0.46	
number of observations	1059	1546	
KS2 English score	**4.12**	**3.91**	0.0000
Standard deviation	0.79	0.93	
number of observations	1027	1502	
KS2 maths score	**4.07**	**3.90**	0.0000
Standard deviation	0.83	0.91	
number of observations	1029	1502	

Notes: this table presents baseline means for a range of variables to enable us to gauge the similarity of the treatment and control groups. Here we present means for the Group 3 workshops plus the control group. Group 3 workshops were those held fortnightly that started at the beginning of the academic year. The last column gives the p-value of a test of equality of means between the two groups: if this is less than 0.1 for a variable the treatment and control groups are said to be significantly different from one another at the 10% level. This means that there is less than a 10% chance that an equally large difference between the two would have arisen by chance if they were in fact the same, i.e. that the two groups are not well matched on this variable at baseline.

Table 8: Descriptive analysis for the three workshop groups separately

Outcome: depression symptoms score

Group 1 treatment: Start Year – Mid Year

	Start Year	Mid Year	End Year	One year later
Treated	8.27	7.71	7.14	7.32
Control	8.03	8.60	7.72	7.22
Gap	0.23	-0.88	-0.58	0.10
SE	*0.37*	*0.44*	*0.42*	*0.39*
p-value	*0.52*	*0.05*	*0.17*	*0.81*
Difference-in-difference		**-1.12**	**-0.81**	**-0.14**
SE		*0.42*	*0.38*	*0.39*
p-value		*0.01*	*0.03*	*0.72*

Group 2 treatment: Mid Year – End Year

	Start Year	Mid Year	End Year	One year later
Treated	10.64	9.25	8.96	8.15
Control	8.03	8.60	7.72	7.22
Gap	2.60	0.65	1.23	0.93
SE	*0.51*	*0.59*	*0.59*	*0.54*
p-value	*0.00*	*0.269*	*0.04*	*0.08*
Average Pre-Policy Gap	1.70			
SE	0.38			
p-value	0.00			
Difference-in-difference			**-0.46**	**-0.77**
SE			*0.50*	*0.52*
p-value			*0.35*	*0.14*

Group 3 treatment: Start Year – End Year

	Start Year	Mid Year	End Year	One year later
Treated	8.31	-	8.74	8.36
Control	8.03		7.72	7.22
Gap	0.27		1.03	1.14
SE	*0.25*		*0.29*	*0.27*
p-value	*0.28*		*0.00*	*0.00*
Difference-in-difference			**0.75**	**0.87**
SE			*0.26*	*0.27*
p-value			*0.00*	*0.00*

Notes: Standard errors (clustered by pupil) in italics. The vertical bar corresponds to the end of treatment so that the period to the right is the post-treatment raw depression symptoms score. Results are comparable with those of Table 29 in Challen et al. (2009), although here a larger sample of pupils is used.

Table 9: Treatment effects for the three experiments pooled
Outcome: depression score
(From specification pooled across all three treatments relative to controls)

	All treatment and control groups							
	Outcome post workshops/end of year (Feb and July 2008)				Outcome at one year follow-up (June 2009)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated*PolicyOn	-0.058	-0.088**	-0.077*	-0.068*	0.033	0.012	0.019	-0.007
	(0.041)	(0.041)	(0.041)	(0.036)	(0.047)	(0.046)	(0.046)	(0.044)
Treated	0.078**	0.119***	0.046		0.078**	0.122***	0.051	
	(0.032)	(0.032)	(0.039)		(0.036)	(0.036)	(0.042)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Number of Pupils	3153	3059	3059	3153	3153	3059	3059	3153
Sample size	7542	7317	7317	7542	6306	6118	6118	6306

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the depression (CDI) score standardised to have a mean of 0 and a standard deviation of 1. Columns (4) and (8) contain controls for mean reversion (5 dummies – initial quintile of depression score*post), and for school trends (school at baseline*post).

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Columns 1-4 present results comparable with those of Table 28 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified slightly. Columns 5-8 present the outcomes for the same pupils one year later.

Table 10: Treatment effects for the three experiments: depression score
(From specification pooled across all three treatments relative to controls)

	Outcome post workshops/end of year (Feb and July 2008)				Outcome at one year follow-up (June 2009)			
Treatment: Start year-Mid year								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated*PolicyOn	-0.103	-0.125*	-0.137*	-0.141***	-0.001	-0.027	-0.037	-0.067
	(0.075)	(0.073)	(0.073)	(0.049)	(0.077)	(0.075)	(0.075)	(0.061)
Treated	0.011	0.038	0.034		0.024	0.047	0.048	
	(0.062)	(0.060)	(0.065)		(0.062)	(0.060)	(0.064)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Treatment: Mid year-End year								
Treated*PolicyOn	-0.195**	-0.198**	-0.175**	-0.029	-0.434***	-0.435***	-0.417***	-0.210**
	(0.087)	(0.084)	(0.085)	(0.069)	(0.125)	(0.122)	(0.123)	(0.091)
Treated	0.182***	0.190***	0.177**		0.427***	0.429***	0.416***	
	(0.064)	(0.063)	(0.072)		(0.099)	(0.096)	(0.106)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Treatment: Start year-End year								
Treated*PolicyOn	0.083*	0.038	0.055	0.021	0.126**	0.101*	0.117**	0.130*
	(0.050)	(0.049)	(0.050)	(0.063)	(0.053)	(0.052)	(0.053)	(0.069)
Treated	0.046	0.107***	0.016		0.036	0.098**	-0.014	
	(0.039)	(0.040)	(0.060)		(0.041)	(0.041)	(0.058)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Sample size	7542	7317	7317	7542	6306	6118	6118	6306
Number of Pupils	3153	3059	3059	3153	3153	3059	3059	3153
p-value of $\chi^2(2)$ test of hypothesis of constant treatment effect	0.0037	0.0135	0.0102	0.0876	0.0001	0.0001	0.0001	0.0104

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the depression (CDI) score standardised to have a mean of 0 and a standard deviation of 1. Columns (4) and (8) contain controls for mean reversion (5 dummies – initial quintile of depression score*post), and for school trends (school at baseline*post). Columns 1-4 present results comparable with those of Table 30 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified. Columns 5-8 present outcomes for the same pupils one year later.

Table 11: Heterogeneity in treatment effects for all three treatments pooled
Outcome: depression score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.068 in 4th column of Table 9)			(Compare to average of -0.007 in final column of Table 9)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
Boys	-0.081* (0.044)			-0.050 (0.054)		
Girls	-0.054 (0.046)			0.030 (0.056)		
p-value from test of equality	0.627			0.229		
Number of pupils Sample size	3153 7542			3153 6306		
Special Educational Needs	-0.205*** (0.060)	-0.216*** (0.077)	-0.175* (0.098)	-0.171** (0.072)	-0.155* (0.093)	-0.167 (0.116)
Not SEN	-0.011 (0.039)	-0.001 (0.056)	-0.032 (0.053)	0.056 (0.048)	0.075 (0.069)	0.017 (0.067)
p-value from test of equality	0.002	0.010	0.159	0.003	0.022	0.125
	3153 7542	1630 3957	1523 3585	3153 6306	1630 3260	1523 3046
Entitled to Free School Meals	-0.113* (0.059)	-0.085 (0.084)	-0.146* (0.083)	0.011 (0.070)	0.009 (0.099)	-0.012 (0.099)
Not FSM	-0.047 (0.040)	-0.063 (0.056)	-0.035 (0.056)	-0.010 (0.050)	-0.008 (0.069)	-0.024 (0.071)
p-value from test of equality	0.300	0.810	0.224	0.780	0.880	0.910
Number of pupils Sample size	3153 7542	1630 3957	1523 3585	3153 6306	1630 3260	1523 3046
KS2 English < Level 4	-0.155** (0.068)	-0.068 (0.088)	-0.281** (0.110)	-0.146* (0.082)	0.005 (0.104)	-0.342*** (0.132)
KS2 English >= Level 4	-0.054 (0.038)	-0.078 (0.054)	-0.040 (0.053)	0.028 (0.047)	0.009 (0.067)	0.025 (0.066)
p-value from test of equality	0.153	0.917	0.033	0.038	0.970	0.006
Number of pupils Sample size	3036 7326	1590 3858	1473 3468	3036 6126	1590 1380	1473 2946
KS2 maths < Level 4	-0.202*** (0.066)	-0.090 (0.094)	-0.309*** (0.092)	-0.153* (0.079)	0.010 (0.113)	-0.318*** (0.110)
KS2 maths >= Level 4	-0.050 (0.039)	-0.080 (0.054)	-0.026 (0.055)	0.018 (0.048)	-0.006 (0.066)	0.035 (0.068)
p-value from test of equality	0.027	0.918	0.003	0.036	0.891	0.002
Number of pupils Sample size	3036 7327	1590 3858	1473 3469	3036 6126	1590 3180	1473 2946

(continued)

Table 11 (continued): Heterogeneity in treatment effects for all three treatments pooled
Outcome: depression score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.068 in 4th column of Table 9)			(Compare to average of -0.007 in final column of Table 9)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
1st quintile baseline depression score	0.015	0.028	-0.005	-0.005	0.091	-0.090
	(0.059)	(0.085)	(0.082)	(0.070)	(0.102)	(0.098)
2nd quintile baseline depression score	-0.054	-0.128	0.021	0.035	-0.036	0.098
	(0.074)	(0.107)	(0.103)	(0.092)	(0.132)	(0.127)
3rd quintile baseline depression score	-0.099*	-0.160*	-0.023	-0.048	-0.043	-0.063
	(0.060)	(0.084)	(0.085)	(0.072)	(0.099)	(0.106)
4th quintile baseline depression score	-0.021	-0.010	-0.051	0.152*	0.121	0.184
	(0.072)	(0.099)	(0.103)	(0.088)	(0.122)	(0.126)
5th quintile baseline depression score	-0.187***	-0.101	-0.321***	-0.121	-0.143	-0.159
	(0.066)	(0.091)	(0.097)	(0.081)	(0.109)	(0.123)
p-value of test of equality 1st-2nd quintile coefficients	0.435	0.224	0.832	0.711	0.420	0.209
p-value of test of equality 2nd-3rd quintile coefficients	0.618	0.804	0.723	0.445	0.962	0.298
p-value of test of equality 3rd-4th quintile coefficients	0.370	0.216	0.829	0.057	0.254	0.108
p-value of test of equality 4th-5th quintile coefficients	0.068	0.468	0.043	0.014	0.080	0.036
p-value of test of equality all quintile coefficients	0.139	0.430	0.048	0.150	0.353	0.182
Number of pupils	3153	1630	1523	3153	1630	1523
Sample size	7542	3957	3585	6306	3260	3046

Notes: Each column-box represents a separate regression. The outcome measure here is the depression (CDI) score, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of depression score*post), and for school-specific trends (school at baseline*post). Columns 1-3 present results comparable with those of Table 31 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified. Columns 4-6 present outcomes for the same pupils one year later.

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 12: Heterogeneity in treatment effects for Group 1 (start-mid year) treatment
Outcome: depression score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.141 in 4th column of Table 10)			(Compare to average of -0.067 in final column of Table 10)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
Boys	-0.109*			0.001		
	(0.065)			(0.078)		
Girls	-0.205***			-0.132		
	(0.069)			(0.083)		
p-value from test of equality	0.2713			0.2008		
Sample size	4907			3838		
Number of pupils	1919			1919		
Special Educational Needs	-0.265***	-0.258**	-0.265*	-0.056	-0.108	0.002
	(0.092)	(0.115)	(0.153)	(0.110)	(0.138)	(0.182)
Not SEN	-0.095*	-0.004	-0.199**	-0.036	0.031	-0.115
	(0.057)	(0.081)	(0.080)	(0.068)	(0.096)	(0.098)
p-value from test of equality	0.0937	0.057	0.6863	0.8703	0.3803	0.5483
Sample size	4907	2635	2272	3838	2030	1808
Number of pupils	1919	1015	904	1919	1015	904
Entitled to Free School Meals	-0.110	-0.048	-0.204*	-0.010	-0.053	0.005
	(0.084)	(0.118)	(0.121)	(0.101)	(0.141)	(0.146)
Not FSM	-0.164***	-0.106	-0.233***	-0.064	-0.012	-0.150
	(0.060)	(0.082)	(0.087)	(0.072)	(0.098)	(0.106)
p-value from test of equality	0.5837	0.6716	0.8353	0.6462	0.8024	0.3634
Sample size	4907	2635	2272	3838	2030	1808
Number of pupils	1919	1015	904	1919	1015	904
KS2 English < Level 4	-0.200*	-0.009	-0.517***	-0.037	0.230	-0.551**
	(0.110)	(0.137)	(0.187)	(0.130)	(0.161)	(0.220)
KS2 English >= Level 4	-0.137**	-0.112	-0.187**	-0.043	-0.075	-0.031
	(0.055)	(0.077)	(0.078)	(0.066)	(0.091)	(0.094)
p-value from test of equality	0.5877	0.4866	0.0874	0.9675	0.0819	0.0226
Sample size	4775	2572	2203	3736	1982	1754
Number of pupils	1868	991	877	1868	991	877
KS2 maths < Level 4	-0.254**	-0.062	-0.468***	-0.211*	0.106	-0.545***
	(0.107)	(0.147)	(0.157)	(0.126)	(0.173)	(0.186)
KS2 maths >= Level 4	-0.136**	-0.107	-0.188**	-0.026	-0.053	-0.025
	(0.056)	(0.077)	(0.081)	(0.066)	(0.091)	(0.097)
p-value from test of equality	0.2989	0.7770	0.0895	0.1669	0.3927	0.0078
Sample size	4772	2568	2204	3732	1978	1754
Number of pupils	1866	989	877	1866	989	877

Table 12 (continued): Heterogeneity in treatment effects for Group 1 (start-mid year) treatment
Outcome: depression score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.141 in 4th column of Table 10)			(Compare to average of -0.067 in final column of Table 10)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
1st quintile baseline depression score	-0.078	-0.019	-0.166	-0.012	0.047	-0.065
	(0.087)	(0.121)	(0.125)	(0.104)	(0.144)	(0.151)
2nd quintile baseline depression score	-0.168	-0.275*	-0.031	-0.078	-0.262	0.157
	(0.111)	(0.156)	(0.158)	(0.133)	(0.187)	(0.192)
3rd quintile baseline depression score	-0.095	-0.113	-0.078	-0.124	-0.043	-0.292*
	(0.093)	(0.124)	(0.140)	(0.111)	(0.146)	(0.171)
4th quintile baseline depression score	-0.008	0.210	-0.228	0.233*	0.378*	0.078
	(0.115)	(0.167)	(0.157)	(0.139)	(0.202)	(0.193)
5th quintile baseline depression score	-0.455***	-0.243*	-0.725***	-0.238*	-0.189	-0.371**
	(0.104)	(0.143)	(0.153)	(0.126)	(0.171)	(0.186)
p-value of test of equality 1st-2nd quintile coefficients	0.5077	0.1772	0.4865	0.6885	0.1743	0.3428
p-value of test of equality 2nd-3rd quintile coefficients	0.5980	0.4029	0.8137	0.7819	0.3409	0.0685
p-value of test of equality 3rd-4th quintile coefficients	0.5433	0.1093	0.4550	0.0366	0.0801	0.1330
p-value of test of equality 4th-5th quintile coefficients	0.0026	0.0328	0.0172	0.0085	0.0264	0.0768
p-value of test of equality all quintile coefficients	0.0164	0.1558	0.0050	0.1049	0.1212	0.1536
Sample size	4907	2635	2272	3838	2030	1808
Number of pupils	1919	1015	904	1919	1015	904

Notes: Each column-box represents a separate regression. The outcome measure here is the depression (CDI) score, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of depression score*post), and for school-specific trends (school at baseline*post). Columns 1-3 present results comparable with those of Table 31 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified slightly. Columns 4-6 present outcomes for the same pupils one year later.

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 13: Descriptive analysis for the three workshop groups separately

Outcome: anxiety symptoms score

Group 1 treatment: Start Year – Mid Year

	Start Year	Mid Year	End Year	One year later
Treated	9.37	8.02	7.08	6.71
Control	8.97	8.25	7.56	6.86
Gap	0.40	-0.23	-0.48	-0.15
SE	*0.37*	*0.42*	*0.39*	*0.37*
p-value	*0.28*	*0.58*	*0.22*	*0.68*
Difference-in-difference		**-0.64**	**-0.88**	**-0.55**
SE		*0.37*	*0.35*	*0.37*
p-value		*0.08*	*0.01*	*0.14*

Group 2 treatment: Mid Year – End Year

	Start Year	Mid Year	End Year	One year later
Treated	11.47	9.41	8.99	8.24
Control	8.97	8.25	7.56	6.86
Gap	2.50	1.15	1.43	1.38
SE	*0.52*	*0.56*	*0.55*	*0.51*
p-value	*0.00*	*0.04*	*0.01*	*0.01*
Average Pre-Policy Gap	1.88			
SE	*0.38*			
p-value	*0.00*			
Difference-in-difference			**-0.44**	**-0.49**
SE			*0.46*	*0.48*
p-value			*0.33*	*0.31*

Group 3 treatment: Start Year – End Year

	Start Year	Mid Year	End Year	One year later
Treated	9.04	-	8.81	8.13
Control	8.97		7.56	6.86
Gap	0.07		1.25	1.27
SE	*0.26*		*0.27*	*0.26*
p-value	*0.80*		*0.00*	*0.00*
Difference-in-difference			**1.18**	**1.21**
SE			*0.23*	*0.25*
p-value			*0.00*	*0.00*

Notes: Standard errors (clustered by pupil) in italics. The vertical bar corresponds to the end of treatment so that the period to the right is the post-treatment raw anxiety symptoms score. Results are comparable with those of Table 32 in Challen et al. (2009), although here a larger sample of pupils is used.

Table 14: Treatment effects for the three experiments pooled
Outcome: anxiety score
(From specification pooled across all three treatments relative to controls)

	All treatment and control groups							
	Outcome post workshops/end of year (Feb and July 2008)				Outcome at one year follow-up (June 2009)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated*PolicyOn	-0.039	-0.055	-0.043	-0.024	0.032	0.020	0.028	0.019
	(0.040)	(0.040)	(0.040)	(0.032)	(0.046)	(0.046)	(0.046)	(0.040)
Treated	0.074**	0.097***	0.037		0.069*	0.095**	0.032	
	(0.033)	(0.033)	(0.039)		(0.036)	(0.037)	(0.043)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Number of Pupils	3125	3036	3036	3125	3125	3036	3036	3125
Sample size	7419	7209	7209	7419	6250	6072	6072	6250

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the anxiety (RCMAS) score standardised to have a mean of 0 and a standard deviation of 1. Columns (4) and (8) contain controls for mean reversion (5 dummies – initial quintile of anxiety score*post), and for school trends (school at baseline*post).

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 15: Treatment effects for the three experiments: anxiety score
(From specification pooled across all three treatments relative to controls)

	Outcome post workshops/end of year (Feb and July 2008)				Outcome at one year follow-up (June 2009)			
Treatment: Start year-Mid year								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated*PolicyOn	-0.120*	-0.139**	-0.156**	-0.103**	-0.092	-0.106	-0.111	-0.050
	(0.071)	(0.070)	(0.070)	(0.043)	(0.077)	(0.077)	(0.077)	(0.055)
Treated	0.025	0.044	0.049		0.037	0.057	0.062	
	(0.060)	(0.060)	(0.064)		(0.061)	(0.061)	(0.065)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Treatment: Mid year-End year								
Treated*PolicyOn	-0.180**	-0.178**	-0.152*	0.030	-0.359***	-0.389***	-0.366***	0.017
	(0.086)	(0.085)	(0.086)	(0.061)	(0.114)	(0.113)	(0.114)	(0.081)
Treated	0.182***	0.193***	0.159**		0.388***	0.402***	0.376***	
	(0.060)	(0.060)	(0.067)		(0.090)	(0.089)	(0.096)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Treatment: Start year-End year								
Treated*PolicyOn	0.122**	0.104**	0.125**	0.063	0.144***	0.134**	0.151***	0.097
	(0.049)	(0.049)	(0.050)	(0.056)	(0.052)	(0.052)	(0.053)	(0.062)
Treated	0.034	0.063	0.011		0.023	0.054	-0.037	
	(0.040)	(0.041)	(0.062)		(0.041)	(0.042)	(0.061)	
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes	No	No	No	Yes
Sample size	7419	7209	7209	7419	6250	6072	6072	6250
Number of Pupils	3125	3036	3036	3125	3125	3036	3036	3125
p-value of $\chi^2(2)$ test of hypothesis of constant treatment effect	0.0003	0.0004	0.0002	0.0288	0.0000	0.0000	0.0000	0.1860

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the anxiety (RCMAS) score standardised to have a mean of 0 and a standard deviation of 1. Columns (4) and (8) contain controls for mean reversion (5 dummies – initial quintile of anxiety score*post), and for school trends (school at baseline*post). Columns 1-4 present results comparable with those of Table 33 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified. Columns 5-8 present outcomes for the same pupils one year later.

Table 16: Heterogeneity in treatment effects for all treatments pooled
Outcome: Anxiety score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.024 in 4th column of Table 14)			(Compare to average of 0.019 in final column of Table 14)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
Boys	-0.069*			-0.019		
	(0.039)			(0.049)		
Girls	0.023			0.049		
	(0.040)			(0.050)		
p-value from test of equality	0.058			0.242		
Number of pupils	3125			3125		
Sample size	7419			6250		
Special Educational Needs	-0.125**	-0.154**	-0.078	-0.088	-0.092	-0.102
	(0.054)	(0.067)	(0.090)	(0.065)	(0.079)	(0.110)
Not SEN	0.016	-0.026	0.041	0.057	0.014	0.077
	(0.034)	(0.047)	(0.050)	(0.043)	(0.058)	(0.063)
p-value from test of equality	0.014	0.078	0.205	0.033	0.207	0.116
Number of pupils	3125	1610	1515	3125	1610	1515
Sample size	7419	3893	3526	6250	3220	3030
Entitled to Free School Meals	-0.121**	-0.208***	-0.059	-0.007	-0.018	-0.036
	(0.052)	(0.072)	(0.076)	(0.063)	(0.084)	(0.094)
Not FSM	0.014	-0.022	0.044	0.033	-0.023	0.077
	(0.035)	(0.047)	(0.053)	(0.044)	(0.058)	(0.067)
p-value from test of equality	0.018	0.017	0.223	0.563	0.954	0.275
Number of pupils	3125	1610	1515	3125	1610	1515
Sample size	7419	3893	3526	6250	3220	3030
KS2 English < Level 4	-0.088	-0.062	-0.097	-0.074	-0.014	-0.166
	(0.061)	(0.076)	(0.104)	(0.074)	(0.088)	(0.129)
KS2 English >= Level 4	-0.020	-0.071	0.018	0.029	-0.036	0.072
	(0.034)	(0.046)	(0.049)	(0.043)	(0.056)	(0.064)
p-value from test of equality	0.280	0.911	0.281	0.174	0.805	0.067
Number of pupils	3041	1572	1469	3041	1572	1469
Sample size	7221	3799	3422	6082	3144	2938
KS2 maths < Level 4	-0.110*	-0.005	-0.203**	-0.056	0.037	-0.148
	(0.058)	(0.079)	(0.084)	(0.071)	(0.095)	(0.105)
KS2 maths >= Level 4	-0.023	-0.087*	0.030	0.021	-0.056	0.080
	(0.034)	(0.046)	(0.051)	(0.043)	(0.056)	(0.065)
p-value from test of equality	0.153	0.320	0.009	0.296	0.338	0.037
Number of pupils	3041	1572	1469	3041	1572	1469
Sample size	7222	3799	3423	6082	3144	2938

Table 16 (continued): Heterogeneity in treatment effects for all treatments pooled
Outcome: Anxiety score
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.024 in 4th column of Table 14)			(Compare to average of 0.019 in final column of Table 14)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
1st quintile baseline anxiety score	-0.006	-0.030	0.006	-0.023	-0.060	0.007
	(0.053)	(0.071)	(0.081)	(0.066)	(0.085)	(0.100)
2nd quintile baseline anxiety score	0.035	-0.070	0.124	0.008	-0.145	0.120
	(0.064)	(0.087)	(0.094)	(0.076)	(0.098)	(0.115)
3rd quintile baseline anxiety score	0.040	0.024	0.048	0.080	0.077	0.073
	(0.057)	(0.076)	(0.088)	(0.070)	(0.091)	(0.107)
4th quintile baseline anxiety score	-0.050	-0.129	0.035	0.085	-0.018	0.169
	(0.058)	(0.080)	(0.084)	(0.071)	(0.095)	(0.105)
5th quintile baseline anxiety score	-0.124**	-0.153*	-0.114	-0.053	0.018	-0.163
	(0.057)	(0.078)	(0.084)	(0.073)	(0.096)	(0.109)
p-value of test of equality 1st-2nd quintile coefficients	0.600	0.705	0.307	0.738	0.481	0.427
p-value of test of equality 2nd-3rd quintile coefficients	0.950	0.382	0.526	0.452	0.071	0.751
p-value of test of equality 3rd-4th quintile coefficients	0.240	0.139	0.912	0.955	0.430	0.494
p-value of test of equality 4th-5th quintile coefficients	0.328	0.819	0.183	0.142	0.776	0.018
p-value of test of equality all quintile coefficients	0.184	0.406	0.348	0.467	0.454	0.157
Number of pupils	3125	1610	1515	3125	1610	1515
Sample size	7419	3893	3526	6250	3220	3030

Notes: Each column-box represents a separate regression. The outcome measure here is the anxiety (RCMAS) score, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of anxiety score*post), and for school-specific trends (school at baseline*post). Columns 1-3 present results comparable with those of Table 34 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified slightly. Columns 4-6 present outcomes for the same pupils one year later.

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 17: Heterogeneity in treatment effects for Group 1 (start-mid year) treatment
Outcome: Anxiety score (RCMAS score)
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.10 in 4th column of Table 15)			(Compare to average of -0.05 in final column of Table 15)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
Boys	-0.090			-0.044		
	(0.058)			(0.070)		
Girls	-0.099			-0.039		
	(0.062)			(0.075)		
p-value from test of equality	0.9057			0.9568		
Sample size	4807			3792		
Number of pupils	1896			1896		
Special Educational Needs	-0.200**	-0.255**	-0.110	-0.136	-0.256**	0.004
	(0.085)	(0.104)	(0.146)	(0.103)	(0.121)	(0.180)
Not SEN	-0.048	-0.042	-0.080	0.007	-0.016	0.013
	(0.051)	(0.069)	(0.075)	(0.062)	(0.080)	(0.094)
p-value from test of equality	0.1071	0.0716	0.8468	0.2063	0.0798	0.9646
Sample size	4807	2578	2229	3792	1996	1796
Number of pupils	1896	998	898	1896	998	898
Entitled to Free School Meals	-0.172**	-0.272***	-0.122	-0.101	-0.239*	-0.020
	(0.076)	(0.105)	(0.113)	(0.093)	(0.123)	(0.141)
Not FSM	-0.053	-0.045	-0.076	0.003	-0.033	0.023
	(0.053)	(0.070)	(0.082)	(0.065)	(0.082)	(0.102)
p-value from test of equality	0.1795	0.0605	0.7288	0.3359	0.1482	0.7929
Sample size	4807	2578	2229	3792	1996	1796
Number of pupils	1896	998	898	1896	998	898
KS2 English < Level 4	-0.143	-0.150	-0.153	-0.139	-0.157	-0.144
	(0.100)	(0.119)	(0.185)	(0.122)	(0.140)	(0.232)
KS2 English >= Level 4	-0.093*	-0.104	-0.103	-0.024	-0.089	0.026
	(0.049)	(0.066)	(0.074)	(0.060)	(0.078)	(0.093)
p-value from test of equality	0.6414	0.7223	0.7979	0.3788	0.6495	0.4814
Sample size	4688	2518	2170	3698	1950	1748
Number of pupils	1849	975	874	1849	975	874
KS2 maths < Level 4	-0.162*	-0.046	-0.292**	-0.199*	-0.099	-0.323*
	(0.097)	(0.128)	(0.148)	(0.119)	(0.150)	(0.187)
KS2 maths >= Level 4	-0.104**	-0.134**	-0.097	-0.018	-0.114	0.059
	(0.049)	(0.065)	(0.075)	(0.060)	(0.077)	(0.094)
p-value from test of equality	0.5712	0.5196	0.2124	0.1490	0.9248	0.0521
Sample size	4685	2514	2171	3694	1946	1748
Number of pupils	1847	973	874	1847	973	874

Table 17 (continued): Heterogeneity in treatment effects for start-mid year treatment
Outcome: Anxiety score (RCMAS score)
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect					
	(Compare to average of -0.10 in 4th column of Table 15)			(Compare to average of -0.05 in final column of Table 15)		
	Outcome post workshops/end of year (Feb and July 2008)			Outcome at one year follow-up (June 2009)		
	All	Boys	Girls	All	Boys	Girls
1st quintile baseline anxiety score	-0.070	-0.067	-0.091	-0.065	-0.073	-0.086
	(0.082)	(0.113)	(0.119)	(0.099)	(0.133)	(0.148)
2nd quintile baseline anxiety score	-0.097	-0.132	-0.088	-0.042	-0.193	0.198
	(0.095)	(0.117)	(0.158)	(0.114)	(0.136)	(0.196)
3rd quintile baseline anxiety score	-0.014	0.034	-0.104	0.202*	0.128	0.240
	(0.096)	(0.126)	(0.148)	(0.116)	(0.149)	(0.181)
4th quintile baseline anxiety score	-0.141	-0.289**	-0.010	0.005	-0.102	0.071
	(0.087)	(0.119)	(0.127)	(0.106)	(0.141)	(0.158)
5th quintile baseline anxiety score	-0.138	-0.081	-0.199	-0.238**	-0.187	-0.302*
	(0.092)	(0.124)	(0.139)	(0.112)	(0.145)	(0.172)
p-value of test of equality 1st-2nd quintile coefficients	0.8227	0.6791	0.9880	0.8756	0.5157	0.2310
p-value of test of equality 2nd-3rd quintile coefficients	0.5229	0.3137	0.9402	0.1199	0.0977	0.8686
p-value of test of equality 3rd-4th quintile coefficients	0.3097	0.0539	0.6165	0.1923	0.2445	0.4625
p-value of test of equality 4th-5th quintile coefficients	0.9778	0.2149	0.2960	0.1024	0.6633	0.0958
p-value of test of equality all quintile coefficients	0.8411	0.3988	0.8944	0.0728	0.4724	0.1288
Sample size	4807	2578	2229	3792	1996	1796
Number of pupils	1896	998	898	1896	998	898

Notes: Each column-box represents a separate regression. The outcome measure here is the anxiety (RCMAS) score, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of anxiety score*post), and for school-specific trends (school at baseline*post). Columns 1-3 present results comparable with those of Table 34 in Challen et al. (2009), although here a larger sample of pupils is used and the specification modified slightly. Columns 4-6 present outcomes for the same pupils one year later.

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 18: Treatment effects for the three experiments pooled
Outcome: absence from school
(From specification pooled across all three treatments relative to controls)

	All treatment and control groups			
	Outcome post workshops/end of year (Sept. 07 - July 08)			
	(1)	(2)	(3)	(4)
Treated*PolicyOn	-0.070	-0.059	-0.048	-0.175***
	(0.048)	(0.046)	(0.044)	(0.053)
Treated	-0.011	0.045	-0.005	
	(0.032)	(0.032)	(0.039)	
Month dummies	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes
Number of Pupils	3810	3744	3744	3810
Sample size	7620	7488	7488	7620

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the fraction of possible school sessions absent over the academic year, standardised to have a mean of 0 and a standard deviation of 1. Column (4) contains controls for mean reversion (5 dummies – initial quintile of absence*post), and for school trends (school at Sept 2007*post, 22 dummies).

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 19: Treatment effects for the three experiments
Outcome: absence from school
(From specification pooled across all three treatments relative to controls)

	Treatment: Start year-Mid year			
	Outcome post workshops/end of year (Feb and July 2008)			
	(1)	(2)	(3)	(4)
Treated*PolicyOn	-0.392***	-0.371***	-0.298***	-0.245***
	(0.076)	(0.075)	(0.073)	(0.074)
Treated	0.197***	0.227***	0.092	
	(0.060)	(0.056)	(0.062)	
Month dummies	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes
	Treatment: Mid year-End year			
	Outcome post workshops/end of year (Feb and July 2008)			
Treated*PolicyOn	-0.238**	-0.212*	-0.049	-0.099
	(0.116)	(0.114)	(0.114)	(0.101)
Treated	0.308***	0.319***	0.031	
	(0.079)	(0.077)	(0.090)	
Month dummies	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes
	Treatment: Start year-End year			
	Outcome post workshops/end of year (Feb and July 2008)			
Treated*PolicyOn	0.079	0.083*	0.043	-0.129
	(0.053)	(0.050)	(0.047)	(0.085)
Treated	-0.146***	-0.076**	-0.035	
	(0.033)	(0.033)	(0.053)	
Month dummies	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	Yes
School Fixed Effects	No	No	Yes	No
Pupil Fixed Effects	No	No	No	Yes
Sample size	7620	7488	7488	7620
Number of Pupils	3810	3744	3744	3810
p-value of $\chi^2(2)$ test of hypothesis of constant treatment effect	0.000	0.000	0.000	0.302

Notes: Standard errors (clustered by pupil) in brackets; the control variables are dummies for gender (1), special educational needs status (1), free school meal status (1), gifted and talented status (1) and Key Stage 2 maths and English performance (10 dummies). The outcome measure here is the fraction of possible school sessions absent over the academic year, standardised to have a mean of 0 and a standard deviation of 1. Column (4) contains controls for mean reversion (5 dummies – initial quintile of absence*post), and for school trends (school at Sept 2007*post, 22 dummies).

Table 20: Heterogeneity in treatment effects for all treatments pooled
Outcome: Absence
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect (Compare to average of -0.175 in final column of Table 18) Outcome post workshops/end of year (Sept. 07 - July 08)		
	All	Boys	Girls
Boys	-0.226***		
	(0.066)		
Girls	-0.115*		
	(0.068)		
p-value from test of equality	0.169		
Number of pupils	3810		
Sample size	7620		
Special Educational Needs	-0.080	-0.233**	0.166
	(0.083)	(0.114)	(0.126)
Not SEN	-0.177***	-0.213**	-0.142*
	(0.058)	(0.089)	(0.076)
p-value from test of equality	0.277	0.876	0.020
Number of pupils	3810	2002	1808
Sample size	7620	4004	3616
Entitled to Free School Meals	-0.196**	-0.333***	-0.048
	(0.081)	(0.120)	(0.109)
Not FSM	-0.150**	-0.191**	-0.103
	(0.060)	(0.088)	(0.081)
p-value from test of equality	0.610	0.287	0.653
Number of pupils	3810	2002	1808
Sample size	7620	4004	3616
KS2 English < Level 4	-0.137	-0.291**	0.144
	(0.092)	(0.124)	(0.139)
KS2 English >= Level 4	-0.143**	-0.180**	-0.116
	(0.056)	(0.084)	(0.075)
p-value from test of equality	0.953	0.396	0.069
Number of pupils	3753	1971	1782
Sample size	7506	3942	3564
KS2 maths < Level 4	-0.114	-0.304**	0.073
	(0.087)	(0.135)	(0.109)
KS2 maths >= Level 4	-0.121**	-0.203**	-0.033
	(0.056)	(0.083)	(0.073)
p-value from test of equality	0.937	0.471	0.360
Number of pupils	3753	1973	1780
Sample size	7506	3946	3560

Table 20 (continued): Heterogeneity in treatment effects for all treatments pooled
Outcome: Absence
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect (Compare to average of -0.175 in final column of Table 18) Outcome post workshops/end of year (Sept. 07 - July 08)		
	All	Boys	Girls
1st quintile baseline absence rate	-0.112	-0.074	-0.135
	(0.087)	(0.129)	(0.116)
2nd quintile baseline absence rate	-0.479	-0.612	-0.285
	(0.441)	(0.566)	(0.907)
3rd quintile baseline absence rate	-0.268***	-0.210	-0.344***
	(0.097)	(0.143)	(0.130)
4th quintile baseline absence rate	-0.220**	-0.343**	-0.112
	(0.098)	(0.146)	(0.132)
5th quintile baseline absence rate	-0.122	-0.464***	0.269**
	(0.097)	(0.141)	(0.132)
p-value of test of equality 1st-2nd quintile coefficients	0.414	0.353	0.870
p-value of test of equality 2nd-3rd quintile coefficients	0.638	0.486	0.949
p-value of test of equality 3rd-4th quintile coefficients	0.711	0.479	0.175
p-value of test of equality 4th-5th quintile coefficients	0.438	0.510	0.029
p-value of test of equality all quintile coefficients	0.632	0.275	0.011
Sample size	3810	2002	1808
Number of pupils	7620	4004	3616

Notes: Each column-box represents a separate regression. The outcome measure here is the fraction of possible school sessions absent over the academic year, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of absence*post), and for school-specific trends (school at Sept 2007*post, 22 dummies).

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

Table 21: Heterogeneity in treatment effects for Group 1 (start-mid year treatment)
Outcome: Absence from school
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect (Compare to average of -0.245 in final column of Table 19) Outcome post workshops/end of year (Sept - July 2008)		
	All	Boys	Girls
Boys	-0.278***		
	(0.098)		
Girls	-0.266**		
	(0.107)		
p-value from test of equality	0.924		
Sample size	4710		
Number of pupils	2355		
Special Educational Needs	-0.207	-0.373**	0.102
	(0.132)	(0.174)	(0.193)
Not SEN	-0.271***	-0.240*	-0.327***
	(0.088)	(0.129)	(0.112)
p-value from test of equality	0.668	0.516	0.042
Sample size	4710	2550	2160
Number of pupils	2355	2375	1080
Entitled to Free School Meals	-0.445***	-0.502***	-0.381**
	(0.123)	(0.175)	(0.161)
Not FSM	-0.141	-0.151	-0.152
	(0.091)	(0.130)	(0.121)
p-value from test of equality	0.038	0.092	0.230
Sample size	4710	2550	2160
Number of pupils	2355	1275	1080
KS2 English < Level 4	-0.329**	-0.518**	-0.008
	(0.155)	(0.206)	(0.225)
KS2 English >= Level 4	-0.222***	-0.200*	-0.279**
	(0.084)	(0.121)	(0.109)
p-value from test of equality	0.521	0.160	0.254
Sample size	4652	2518	2134
Number of pupils	2326	1259	1067
KS2 maths < Level 4	-0.535***	-0.669***	-0.412**
	(0.149)	(0.229)	(0.191)
KS2 maths >= Level 4	-0.154*	-0.175	-0.151
	(0.082)	(0.120)	(0.111)
p-value from test of equality	0.017	0.045	0.203
Sample size	4648	2518	2130
Number of pupils	2324	1259	1065

Table 21 (continued): Heterogeneity in treatment effects for start-mid year treatment
Outcome: Absence
(From separate specifications relative to controls)

	Estimated Treatment*Policy Effect		
	(Compare to average of -0.245 in final column of Table 19)		
	Outcome post workshops/end of year (Sept - July 2008)		
	All	Boys	Girls
1st quintile baseline absence rate	-0.320**	-0.372*	-0.258
	(0.147)	(0.214)	(0.186)
2nd quintile baseline absence rate	-0.574	-0.658	-0.423
	(0.712)	(0.860)	(1.339)
3rd quintile baseline absence rate	-0.278*	-0.249	-0.363**
	(0.150)	(0.232)	(0.180)
4th quintile baseline absence rate	-0.207	-0.100	-0.336*
	(0.144)	(0.205)	(0.187)
5th quintile baseline absence rate	-0.278**	-0.478**	0.032
	(0.138)	(0.186)	(0.197)
p-value of test of equality 1st-2nd quintile coefficients	0.726	0.747	0.903
p-value of test of equality 2nd-3rd quintile coefficients	0.683	0.644	0.964
p-value of test of equality 3rd-4th quintile coefficients	0.717	0.617	0.911
p-value of test of equality 4th-5th quintile coefficients	0.708	0.156	0.154
p-value of test of equality all quintile coefficients	0.972	0.675	0.578
Sample size	4710	2550	2160
Number of pupils	2355	1275	1080

Notes: Each column-box represents a separate regression. The outcome measure here is the fraction of possible school sessions absent over the academic year, standardised to have a mean of 0 and a standard deviation of 1. Regressions include controls for mean reversion (5 dummies – initial quintile of absence*post), and for school-specific trends (school at Sept 2007*post, 22 dummies).

Here, and in subsequent tables, the level of significance of a coefficient is indicated by asterisks after it: one asterisk means that it is significant at the 10% level (p<=0.1 – there is a 10% chance that this could have arisen by chance); two asterisks that it is significant at 5% (p<=0.05); and three that it is significant at 1% (p<=0.01).

4. Qualitative Findings

Summary of Qualitative Findings

The quantitative findings reported earlier all relate to the first cohort of Year 7 pupils who attended UKRP sessions in 2007-08 and were subsequently followed up. 2007-08 was the first year in which the programme was delivered. In this section we report findings from the qualitative element of the research project. Here we focus on how the UKRP had been implemented in schools and how the programme was operating in schools in its third year of delivery (2009-10). It is important therefore to regard the quantitative and qualitative elements of the evaluation as discrete elements of the research project.

Ten UKRP schools were visited in 2007-08 to collect qualitative interview data to supplement the quantitative analysis that forms the core of the UKRP evaluation. Findings from these visits were reported in the previous interim report (Challen et al., 2009). Nine of the ten case study schools were visited again in the autumn term of 2009-10 to examine how the implementation of the UKRP had progressed in those schools.

In 2009-10 the UKRP was being delivered to 100% of the Year 7 cohort in seven of the nine case study schools visited. At one of the other schools a decision had been taken to discontinue delivery of the programme and at the other delivery in 2009-10 had been postponed perhaps indefinitely.

In three schools the UKRP was being delivered primarily by teachers and in the other four schools it was delivered primarily by non-teachers (for example, teaching assistants, cover supervisors or learning mentors). The schools in which the UKRP was delivered primarily by teachers all had impressive track records in promoting pupils' academic progress. One interpretation would be that this may have enabled teachers to focus on pupils' well-being, through the UKRP, rather than focusing more exclusively on attainment.

At schools where more non-teachers delivered the UKRP it appeared that pay and holiday arrangements, workloads, and non-teachers seeing the UKRP as a good career development opportunity may have contributed to this drift to non-teachers.

Senior management backing for the UKRP was clearly very important to its successful implementation and could vary substantially. Management backing could be shown through financial support and through giving the UKRP relatively high priority when deciding on the school timetable.

The UKRP was accommodated in the curriculum either as a stand along subject, timetabled once per fortnight or timetabled weekly in conjunction with another subject. For example, in some schools the UKRP was delivered during English lessons, PSHE lessons or as part of a PLTS programme.

Interviewees were generally positive about their experience of the programme. Evidence from interviews indicated that the emotional content of UKRP sessions could vary substantially. Pupils could sometimes raise upsetting issues during UKRP sessions. In keeping with the previous report (Challen et al., 2009) interviewees spoke positively about the quality of UKRP training though some expressed reservations about the quality of the teaching materials.

Introduction

As part of the UKRP evaluation, the DCSF requested that return visits be made to the case study schools that were visited in 2007-2008 (reported in Challen et al., 2009). At that time, visits were made to ten of the twenty two secondary schools participating in the UKRP. Case study schools were selected to reflect variation in the proportion of the Year 7 cohort receiving UKRP sessions in the first year of the programme. They were also selected to reflect variation in levels of pupil attainment, the level of eligibility for free school meals and school Contextual Value Added (CVA) scores. In the first year of the programme the UKRP was delivered by a mixture of teachers, school support staff and by facilitators not employed by individual schools (most often local authority employees). Case study schools were also therefore selected to reflect this variation. In addition, at least three schools were selected from each of the three participating local authorities.

The return visits to the case study schools were undertaken in the autumn term of 2009-10. It was possible to revisit nine out of the ten original case study schools. At each school interviews were requested with up to four members of staff including a senior manager, two UKRP facilitators (one trained in the original 2007 cohort and one trained at a subsequent training event) and, if the UKRP was linked with any other programmes within the school (such as Personal, Social and Health Education (PSHE)) a member of staff responsible for that programme.[35]

In total, interviews were carried out with thirty-one members of staff at case study schools, all of which were fully transcribed. These were supplemented by five telephone interviews carried out with interviewees unavailable during case study visits (including UKRP co-ordinators for two of the three participating local authorities).

In what follows we describe how the UKRP has developed within the case study schools and, as a means of summarising that development, we split the case study schools into three groups. We then go on to discuss some of the most prominent organisational issues raised during the case study visits, namely a drift, in some schools, towards the programme being delivered by school support staff, the importance of management backing for the programme and finding a place for the UKRP within the school curriculum. We then go on to discuss some of the views expressed by interviewees, including their overall views on the UKRP and focusing in particular on two important issues raised: dealing with emotive issues during UKRP sessions; and facilitators' views on the UKRP course materials. We then draw together some of the findings from the case study visits in the concluding section.

The case study schools in 2009-10

As we reported in the previous interim report, the UKRP was implemented within schools in a variety of ways. For example, some schools had undertaken substantial forward planning for the UKRP and it slotted easily into the school curriculum and timetable while at others the process of identifying opportunities to deliver the programme was undertaken after the initial

[35] In 2009, interviews were carried out with pupils following the UKRP in order to provide examples of pupils' recall and reported use of the UKRP skills they had learned during sessions. No interviews with pupils were carried out when schools were revisited.

cohort of facilitators returned from their training. In addition, we reported on the varying ways in which the programme was timetabled, depending on the other constraints faced by the schools, and the different curriculum homes found for the programme.

The scale of the UKRP at these nine schools varied at the time of the first case study visits in 2007. At that time, at three of the schools the UKRP was being delivered to the whole Year 7 cohort, at four of the schools at least one third of the Year 7 cohort was receiving the UKRP and at two schools no more than 25% had followed a UKRP course. Two years later the picture had changed substantially. At seven of the nine schools the UKRP was being delivered to 100% of Year 7 pupils (although one of these schools did not deliver the UKRP to any pupils in 2008-09). Of the remaining two schools, at one a decision had been taken to discontinue the programme. At the other school delivery of the UKRP was reported to have been postponed and it appeared likely that the UKRP would not be delivered at all in 2009-10.

Of course the continuation, and in some schools expansion, of the UKRP within the case study schools is probably the best indicator of the schools' commitment to and satisfaction with the UKRP.

In the first instance, for ease of presentation, we split the schools into three groups. The first group of three schools comprise those in which the UKRP persists as a predominantly teacher-led programme. In the second group of three schools the UKRP was embedded in the curriculum and the majority of UKRP facilitators were auxiliary staff. In the third group the UKRP was not been delivered to pupils in all three years since the programme's launch in 2007-08. Three case study schools have been placed in each of these three categories and some key information about the schools is shown in Table 22.

Table 22: Key information about the nine case study schools revisited in 2009-10

Groups of schools	Trained facilitators at the school	Size of Year 7 UKRP Cohort (all pupils unless stated)	Teaching slot	Contextual information	
				CVA scores, 2007-09‡	Most recent OFSTED judgement
1. Teacher- led	17		English	++	Satisfactory
	10		UKRP	++	Outstanding
	8		PSHE	++	Good
2. Embedded, support staff-led	10		PSHE	+	Satisfactory
	8		PLTS*	-	Satisfactory
	7		UKRP	--	Inadequate
3. UKRP not delivered in all three years	6	Two classes (start postponed)	PSHE	++	Outstanding
	3		UKRP	++	Good
	3	Discontinued	Citizenship	--	Satisfactory

*Personal Learning and Thinking Skills

‡Symbols refer to statistically significant CVA scores in the previous three years. + means one statistically significant positive CVA score in the previous three years, ++ means two such scores etc. CVA scores are published as part of the DCSF's Achievement and Attainment Tables. They focus on pupils' attainment at Key Stage 4, taking account of their prior attainment at Key Stage 2 and pupil and school characteristics that are beyond the control of the school. They therefore provide the best publicly available indicator of a school's effectiveness in promoting academic attainment. A 'statistically significant CVA' score refers to a school for which the CVA confidence interval does not include 1000 (the mean predicted score for pupils at that school). A full explanation of DCSF CVA scores is available from the DCSF performance tables website.

UKRP operates as a teacher-led programme (group 1)

At three of the case study schools teachers, as distinct from support staff, were continuing to deliver most of the UKRP sessions in 2009-10. At these schools there had been continuing demand from staff to undertake the UKRP training. In one of the schools seventeen staff had undertaken UKRP training, of whom the majority were teachers. In the other two schools at least ten staff had received UKRP training and again the majority of those staff were teachers rather than auxiliary staff (at one of the schools two of the trained staff had since left the school).

In all three schools the programme was securely embedded in the curriculum. For example, at one the UKRP had its own, named lesson on the Year 7 timetable, in another the UKRP was delivered by members of the English department with one English lesson per week

devoted to the UKRP for part of the year, and in the third school the UKRP was delivered as part of a Year 7 PSHE programme which also incorporated some elements of the SEAL programme.

For one of the schools the organisational arrangements for the UKRP had been established in the first year of the programme and it had been delivered on the same basis for the subsequent two years. This was however not the case for the other two schools.

One had made a tentative start in the first year of the programme, delivering the UKRP to a small proportion of the Year 7 cohort during PSHE lessons. The programme had been extended in the second year but delivery had been limited by the fact that every class in the school undertook PSHE at the same time and therefore the proportion of the year group that could follow the course was constrained by the number of facilitators trained. In the third year, at the instigation of the headteacher, the entire English department received UKRP training (10 teachers), thus enabling the programme to be delivered to the whole Year 7 cohort thereafter.

At the third school in this group, the programme was delivered to the whole Year 7 cohort from the first year of operation although the organisation of that delivery changed over the three years. While the UKRP was initially delivered in fortnightly sessions, in the third year of the programme it was incorporated into a new PSHE programme also including elements of the SEAL curriculum, and consequently the UKRP was then delivered at weekly sessions with Year 7 pupils. The school had successfully maintained delivery to the whole cohort despite two out of three of the teachers trained in the first year then leaving the school.

In the three schools, senior managers and UKRP co-ordinators were knowledgeable and clearly understood how the UKRP fitted in to the organisation of the school and had clear ideas about future delivery of the UKRP and how to ensure the continuing success of the programme.

Interestingly, in all three schools, in the last three years, the schools had achieved a statistically significant positive CVA score from Key Stage 2 to GCSE in two of the three years. In their most recent OFSTED inspections, one school had been graded 'outstanding' and one graded 'good'. The third school, having previously been graded 'outstanding', had received a 'satisfactory' grading in its very recent OFSTED inspection. It is however worth noting that interviewees (from both within and outside the school) attributed this result to the new OFSTED framework which gives greater emphasis to attainment levels. The intake of the school was traditionally very disadvantaged and attainment levels correspondingly low and close to the National Challenge threshold.

UKRP is embedded in the school curriculum and delivered by auxiliary staff (group 2)

At three of the case study schools the staff trained to deliver the UKRP were predominantly non-teachers such as higher level teaching assistants, learning support assistants, learning mentors and cover supervisors. At all three schools between five and ten members of staff had received UKRP training. At each, the UKRP was delivered to the whole Year 7 cohort.

In one of the schools the UKRP was timetabled as a named lesson, delivered to pupils fortnightly. In the other two schools UKRP sessions were delivered weekly, comprising part of a PSHE course and a Personal Learning and Thinking Skills (PLTS) programme. While the delivery of the programme by auxiliary staff may not in itself be indicative that the programme was less secure than in the group 1 schools, in each school there were possibly signs that the programme may be less secure than in the group 1 schools (for example, a lack of knowledge of the programme from senior staff responsible or recent changes to senior management or an on-going curriculum review). Of course, it should be noted that the appearance of firm foundations for the programme at the time of the first case study visit was not necessarily predictive of how the UKRP was faring two years later. Similarly, it would be unwise to make predictions about the future operation of the UKRP within the case study schools.

One of the schools in this group had achieved a statistically significant positive CVA score in one of the previous three years. The other two schools had both achieved statistically significant negative CVA scores, with one of the schools receiving this score in two of the three previous years. The most recent OFSTED reports for the three schools had graded two 'satisfactory' and one 'inadequate'.

UKRP was not delivered in all three years (group 3)

In three of the case study schools the UKRP had failed to thrive during its first three years for a disparate set of reasons. At all three schools a lack of management backing had been a factor in this situation.

One senior manager acknowledged that the school had not fully committed itself to the programme from the outset and indeed it had only been delivered to a small number of groups in the first year. The school had also suffered particularly badly from staffing difficulties. Both of the prospective trainees for the second UKRP training event had pulled out at the last moment. In addition, two of the three staff trained in 2007 were not working at the school in 2009-10. As a result of these staffing problems, the programme had not been delivered to any pupils in its second year. However, after recruiting two members of the school's auxiliary staff to be trained in the third cohort of trainees (and with the assistance of a facilitator employed by the local authority), the programme was being delivered to all Year 7 pupils in the third year of the programme. Nevertheless, the programme appeared less secure than in any of the group 1 or group 2 schools.

At another group 3 case study school UKRP provision had been discontinued part way through the second year (2008-09). Only one teacher from the school had been trained in the initial cohort of trainees, with two members of the support staff trained in the second year. Despite encouraging early signs for the UKRP, delivery was discontinued for combination of (numerous) reasons. These were reported to include a shortage of appropriate teaching spaces, behaviour management problems in the UKRP groups, the departure of the headteacher and senior manager responsible for the programme, and the school being identified as a National Challenge school precipitating a shift in resources from the UKRP to improving levels of attainment.

In the third school, the initial commitment to the UKRP had been very tentative with the senior manager responsible commenting at that time that the programme did not have the backing of the senior management team as a whole. The programme was delivered to a very small number of groups in the first year with only a slight expansion in the second year. Delivery of the UKRP was reported to have been delayed in the third year (and at the time of the case study visit to the school). The reported delay followed a debate as to the form of consent required from parents for their children to participate in the programme. It was seen as particularly important to gain positive parental consent (rather than giving them an opportunity to opt out) because of an imminent OFSTED inspection. The senior manager responsible for the UKRP subsequently suggested that, despite a personal belief in the value of the programme, it was unlikely that the UKRP would be delivered at all that year.

Interestingly, while one of the schools in this group was identified as a National Challenge school and had received a statistically significant negative CVA score in two of the previous three years, both of the other schools had relatively high levels of attainment. Each of the other schools had achieved statistically significant positive CVA scores in two of the previous three years. The most recent OFSTED reports for the schools had judged them to be 'satisfactory', 'good' and 'outstanding'.

Thus we see in the nine case study schools revisited in 2009-10, the UKRP had developed in a wide variety of ways. In the three 'group 1' schools the programme had become firmly established and was in most cases delivered by teachers. Interestingly, all three of these schools had impressive track records in pupils' achievement and (with one exception) OFSTED judgments. Group 2 schools were perhaps under greater pressure to improve pupils' academic achievements and delivery of the UKRP had largely shifted to auxiliary staff rather than teachers. Group 3 schools comprised three disparate schools in which the UKRP had failed to thrive for a wide range of reported reasons.

A drift to non-teachers?

As we have seen, the three schools in which the programme continued to be predominantly teacher-led each had a record of being a value adding school in two of the three previous years.

In the other six schools (group 2 and group 3) there was certainly a drift towards the programme being delivered by non-teachers. To express this numerically, in the first cohort of trainees who attended training in Philadelphia and who came from these six schools, eleven trainees were teachers and six were support staff. In the subsequent two years however trainees from these schools comprised four teachers and nineteen support staff.

Several interviewees referred to this tendency, for example:

> initially, [recruitment] was aimed at heads of year and deputy heads of year… [teachers] more in a pastoral role... And then after that… it was maybe more geared towards having more… learning support staff, who generally deal with… the lower ability students… [and] children with emotional and behavioural [problems] so they

> might [be] staff who have a different relationship with the kids. [teacher, support staff-led school, group 2 school]

Two of the original cohort of trainees at other schools identified a link between this shift from having teachers deliver the UKRP and pressure to achieve key attainment targets. One commented:

> it's very difficult for teachers because as you can see, we've kind of been pushed out of UKRP because our timetables are, the commitments are quite big elsewhere in terms of exam results and things like that, especially core subjects like myself. So I would say that… it's all non-teaching staff who are doing it now. [teacher, group 2 school]

This remark was echoed by a teacher in another school (group 3):

> And the senior management team… decided… I was having to spend too much of my time with UKRP and not enough time with English and so the decision was made, strategically, to stop teaching UKRP myself… Because there are so many pressures on English, so many pressures on English results. [teacher, group 3 school]

Importantly, two of the teachers quoted above taught in schools with examination results close to the National Challenge threshold.

However, some senior managers suggested that the drift towards delivery by non-teachers was not the result of deliberate policy decisions. For example, several explained that there was an open call for members of staff to apply to attend the UKRP training, but as one senior manager stated:

> one of the areas we've really struggled on… is… recruiting teaching staff to do the training to deliver the programme… It was actually difficult getting people on board, from the teaching staff specifically (senior manager, group 3)

Several reasons were identified by interviewees for the drift from teachers to support staff. First, several referred to the difficulty of having to persuade applicants to attend training during their summer holiday, and this was particularly so among staff at group 2 and 3 schools. For example:

> *Do people want to do [the training]?*
> …no, they don't. You're turning round to them and going 'Give us one week of your holidays', that's where the problem comes in, you know? (teacher, group 3)

> it was a bit hard to find a member of staff who was willing to give up [some] of their summer holiday, basically (teacher group 3)

The same point was made by a facilitator from a group 2 school's support staff:

Is there anything that you'd like to tell me about it that I haven't asked you about and that I should have?

Just if they're going to [do] another training course, it might be easier just in term time, really, not ask people to give up a week of their holiday. (auxiliary staff, group 2)

Similarly this point was echoed by a teacher who had been approached with a view to undertaking a higher level of training in order to assist with the training of future UKRP facilitators:

> I was sent a very complimentary email, it was very nice… they contacted me to be a trainer last summer, which I declined because it would have effectively meant giving up 2 weeks of my summer holiday… Obviously… I know that it would benefit the kids… but at the end of the day, it was 2 weeks of my holiday. (teacher group 2)

While the loss of holiday time was clearly a disincentive to participate in the programme, it is important to note that training places were nevertheless filled, and sometimes by offering additional training places to schools at which there were higher levels of demand (exemplified by the three group 1 schools).

Interviewees also gave some insights into why there may have been a greater incentive to participate in training for non-teachers than teachers. As one senior manager stated in relation to teachers:

> you're effectively asking people to do something in addition to what they currently do (senior manager, group 3 school)

That is, teachers embarking on the programme would then be committed to the greater preparation time associated with new courses. However, in several of the schools non-teachers were allocated preparation time specifically for the UKRP. In addition, several senior managers pointed out that for non-teachers the UKRP was seen as an attractive career development opportunity, for example:

> I think… the other TAs [Teaching Assistants] got wind it's a really good training opportunity as well… So, you know, more became interested in it as well. And a lot of them have used it as a springboard and gone on to do or they're looking to do PGCE [Post Graduate Certificate in Education], you know, anyway so this is of an ideal stepping stone from teaching assistant to teacher. (senior manager, group 2 school)

Indeed, this was illustrated by the responses of a number of non-teacher interviewees who explained that the programme fitted with their career plans to become teachers. In addition, the senior manager quoted above went on to explain that in the school, as support staff were on term-time only contracts, the school had also paid the teaching assistants and cover assistant to attend the training whereas this was not the case for the teaching staff who had attended.

From the point of view of senior managers, delivery of the UKRP by non-teachers may offer two firm advantages over a teacher-led programme. First, the arrangement offers much greater timetabling flexibility. Second, as we have seen, other demands placed on schools,

such as a drive to meet attainment targets, could place teachers under pressure and potentially place a teacher-led programme in a vulnerable position. On the other hand, it was suggested during the first round of case study visits that the involvement of teachers, and especially senior teachers, can enhance the status of the programme within a school.

In contrast to the schools in which non-teachers had taken an increasingly prominent role in delivering the programme, interviewees from the three schools that continued to run 'teacher-led' UKRP arrangements did not report any recruitment difficulties. For example, one stated (to the evident surprise of the interviewer):

> I don't think there's any shortage of interest in delivering the programme.
> *Do they still have to give up a week of the summer holiday?*
> Yes, yeah just the way it was scheduled yeah.
> *But there was interest despite that?*
> Yeah absolutely yeah.
> *Yeah and no need to… make any compensatory arrangements?*
> No not that I'm aware of I think there was, I think 3 teaching staff this year? And 2 non teaching staff.
> (Teacher, group 1 school)

In one of the schools with a teacher-led programme, senior staff approached particular members of staff to invite them to train as facilitators. The UKRP co-ordinator at that school expressed unhappiness at the fact that UKRP training participants from some schools were paid to attend while those from others were not. Indeed, this interviewee also expressed the view that it was useful that participants had to give up some of their holiday time as this was a means of demonstrating their commitment to the programme.

Thus we have seen that there was a drift towards the programme being delivered by non-teachers in several of the case study schools and that interview data suggests that there may have been several reasons why this may have been the case including the pressure of attainment targets, teachers' workload, holiday and pay arrangements, and the perception of the UKRP as a good career development opportunity for non-teachers.

There was no agreement among interviewees as to whether this drift should be a cause for concern. On the one hand, some teachers expressed the view that the programme should only be delivered by qualified teachers. And in addition, managers' descriptions of the organisation of the programme sometimes implied a lack of confidence in the capacity of at least some non-teachers to deliver the programme. For example:

> We're very, very cautious… sometimes even the teaching assistants and pastoral managers are not comfortable with a class on their own, therefore you're putting two people in [to a UKRP session to compensate for that] (manager, group 3 school)

Similarly at a group 2 school the senior manager explained that class teachers were always available (though out of the room) to support non-teacher UKRP facilitators in dealing with behaviour management problems – 'we don't just leave them'.

Non-teachers frequently acknowledged how daunted they had been at the prospect of leading UKRP sessions with groups of up to 15 pupils. For example:

How well do you think the training prepared you for delivering the Resilience Programme?
It – and I'm only a learning support assistant, so I'm not a teacher – so it was very, very daunting when I, on my way back with these three massive A4 folders, that I've got to sit down and try and work out a lesson. I didn't know what an hour's teaching lesson would mean, as I'm sitting there… and trying to prepare these lessons. So that was very, very scary. That was very, very daunting. (non-teacher, group 2 school)

However, they often went on to describe how their fears had been allayed when they began to deliver the course. Indeed, when senior managers were asked directly about whether they believed it was better for teachers or non-teachers to deliver the programme, they tended to emphasise the importance of the characteristics and interests of potential trainees rather than their job title.

One means of illustrating variations in facilitators' ability to identify and communicate clearly about key aspects of the programme through interview data was to ask how they would describe the main elements of the UKRP to someone unfamiliar with the programme. However, while a wide variety of descriptions were provided, they did not suggest that there may be a systematic difference between teachers' and non-teachers' abilities to communicate the content of the programme. For example, some interviewees provided brief descriptions that drew attention to key messages from the programme relating to cognition:

I would emphasise the positive thinking side of things. Trying to get kids to see things in a different way. Trying to get kids to make the link between events and how that makes them feel. And how they can actually influence that, how they can actually use a skill to influence the way they feel… (non-teacher, group 1 school)

The main thing is… trying to give them skills so … that students are actually able to think in a more flexible and accurate way… They don't just stick with one rigid thinking style, particularly a pessimistic one. So they're able to use the different skills to be more flexible in their thinking and actually be more accurate; actually find evidence for their problems and different solutions for their problems… (teacher, group 3)

On the other hand other interviewees provided much less precise or arguably more partial summaries of the programme:

Obviously about problem solving, giving students the opportunity to talk about things in their lives, good or bad really… - that's it really. Just giving students the opportunity to talk about themselves and their lives; and working out strategies to make things better. (non-teacher, group 3 school)

It's giving children the skills to be able to deal with situations and not to use violence as their first port of call, in my opinion. It is very much about giving them the skills to be able to look after themselves, but also to keep them out of trouble. And also to realise that there are different ways of looking at things; that not everything's completely black and white, there is sometimes that middle ground that we just forget about. (teacher, group 1)

Perhaps, rather than focusing on categories of staff, it is most important that schools strive to recruit the most suitable staff to become facilitators, therefore drawing them from the widest possible pool.

The importance of senior management backing for the programme

At the school with a teacher-led programme in which the UKRP was delivered during English lessons, the recruitment strategy had been especially directive. As was noted earlier, the two years in which the programme had been delivered during PSHE lessons at that school had been problematic:

> it was never tight enough… you couldn't get the staff who were trained because they were already delivering PSHE to their form groups (senior manager, group 1)

The senior manager went on to explain that the headteacher had suggested that the entire English faculty should be trained to deliver the programme in future. As the senior manager continued:

> the English faculty were already very good at talking about empathy, and empathetic skills… and some of the subjects that you have to be able to talk about and deliver through Penn are already being done through the English curriculum. The staff were very comfortable with it. (senior manager, group 1)

The head of faculty's initial response to the suggestion was positive:

> I thought it was an excellent idea… I think your initial concern as head of faculty is, my faculty works very hard all year… obviously there is a lot of pressure on English as a core subject. And my initial concern was the fact that we were taking a week's holiday off staff… But as usual, they all stepped up to the mark…

The head of faculty went on to explain:

> I said to people… obviously, it's not compulsory, it's the holiday, but we're going to deliver this through English next year. If you want to come we will pay you, and everybody came, in fact, even staff who were joining us this year and hadn't taught here previously came. (head of faculty, group 1)

Both the head of faculty and the headteacher had then attended the training programme, taking the view that it was important to lead by example (although the headteacher had not subsequently been called on to deliver the course). This school perhaps best exemplifies the importance of management backing to facilitate the successful operation of the programme. In other respects too this school could be seen as providing a Rolls Royce implementation model.

Having trained all English teachers to deliver the programme the school was confident that it would be in a position to continue offering the UKRP, should it choose to do so, to all Year 7 students in subsequent years, regardless of the exigencies of the school timetable. In addition, it was decided that the UKRP should be delivered simultaneously to all pupils during the same English lesson (chosen to maximise the number of other UKRP-trained staff

available to teach – as classes were split in half). It was pointed out that by teaching the UKRP to all UKRP groups during the same English lesson, it was then possible for pupils to move from one English set to another without needing to change the membership of UKRP groups. The arrangements also ensured that there was plenty of time in which to deliver the UKRP should it require more lessons than anticipated.

While we might view this as a Rolls Royce delivery model it is of course important to keep in mind that this model was in its first year of operation and, as at the other schools, should the pressure to improve levels of achievement and attainment increase, pressure on the UKRP may also increase.

A clear contrast may be drawn between this school and two of the group 3 schools to illustrate the importance of management backing. In the first year of its operation at one school, staff had been frustrated that the senior manager responsible for the timetable had been unwilling to accommodate the constraints that would be imposed by the UKRP (namely splitting classes in half and the programme being delivered by particular UKRP-trained staff). Two years later an interviewee stated:

> We do need more teachers to be trained… if you have your teaching staff who are trained to do it, then you can actually say to your headteacher, 'look, I'd like to do it in this… lesson… and my team are trained' (teacher, group 3)

That is, rather than having the timetable designed to accommodate the programme or senior managers directing staff to be trained, the teacher envisaged staff being trained first in order to persuade the headteacher that the programme could be accommodated within the existing timetable. While part of the implementation problem at this school arose from its timetabling, other problems stemmed from the fact that it was felt that the auxiliary staff who had received UKRP training did not have the confidence or experience to deliver the UKRP alone, thus also limiting the number of groups that could potentially follow the programme. As the senior manager responsible for the UKRP put it:

> Last year [2008-09] seemed to really just muddle along as it had done in the previous year really. And as I explained to you at the time, it really was "muddle along". And since September [this year] there's been nothing. (senior manager, group 3)

Given the lack of management backing that the programme had received from the outset, and its somewhat limited implementation, it is perhaps unsurprising that the programme had been allowed to stall as a result of what, at some schools may have been viewed as a relatively minor problem. This problem arose from the new OFSTED framework and an imminent inspection, as the senior manager explained:

> As a result [of the new framework in relation to safeguarding] all schools were looking at their permissions letters… the UKRP one was 'if you don't reply we assume it's OK [for your child to attend the UKRP sessions]'… It was felt… that that was not the appropriate thing for the UKRP. And consequently it was suggested that we would have to get [parents] to opt in to it. And [staff] weren't comfortable with that and so that just didn't happen. And as a consequence, none of the sessions has been delivered… We were then OFSTED-ed and found to be absolutely super duper in terms of our safeguarding. (Senior manager, group 3)

The outcome of the inspection overall was that the school was judged to be 'outstanding'. Given the school's good performance in CVA analysis also, senior managers will doubtless feel vindicated in the priorities pursued within the school. The implementation of the UKRP had not however been a priority.

While the senior manager interviewed expressed support for the potential value of the UKRP and frustration that it had not been successfully implemented at the school, it appeared likely that the programme would be discontinued. This possible demise of the UKRP within the school is however tinged with an irony that was not lost on the senior manager responsible. When asked in what way the school believed OFSTED would view consent letters fitting in with safeguarding, the senior manager replied:

> Because of the content of the course it's felt that whenever parents' permission for [their children] to be involved is sought, it's felt that… they need to opt in to it… And we are a school, and the type of children that were chosen for the UKRP were the type of children, who wouldn't get the written consent letters very easily. It would be a huge chasing admin job.
> *Who is being safeguarded in that case?*
> The children from the content. From having to deal with some of the social and emotional issues that it's felt that the content will bring up –
> *So* –
> And that's the very purpose why we want to deliver it. [laughs]… I find it just one of those things that just defeats the object. (Senior manager, group 3)

It is perhaps also informative to consider the case of the other group 3 school in which the programme had been discontinued. The removal of management support for the scheme (as a result of the headteacher leaving the school) was one component of the demise of the UKRP, although only one of several.

The move to a new building, a difficult year group, inexperienced facilitators, the departure of the headteacher, promotion of the original facilitator and the school becoming a National Challenge school were all identified as contributory factors. In the first year of the UKRP, the senior manager responsible for the UKRP, who also taught citizenship to Year 7 classes, had arranged for those classes to be split in half so that half could receive the UKRP while the other half would be taught citizenship in a small group. The UKRP teacher explained that the first year for the programme had been seen as a success, in the sense that students appeared to have enjoyed the programme and some pupils could give examples of when they had used the UKRP skills they had learned in real life.

The UKRP teacher was initially the only trained facilitator at the school and two auxiliary staff had undertaken the UKRP training in the summer of 2008. At the same time the school was identified as a National Challenge school and also moved into a new, largely open plan, building, the teacher explained:

> Lots of the classrooms don't have a door, they have half a wall… onto the corridor. Which is fine for when you're doing… normal English teaching… - obviously problems with sight lines and kids in the corridor, levels of volume and things… [But] when it comes to confidentiality… 'what's said in this room stays in this room'. Well,

> [if] the room physically doesn't have a wall, then that kind of negates the whole thing. So we had quite a lot of logistical issues finding rooms which had doors… [And] although some of the classes that I had started to… gel, others didn't. Because I was being asked to teach it in an open plan room. And that was very, very problematic. (teacher, group 3)

In addition, the two members of the school support staff trained to run UKRP groups encountered behaviour management problems:

> I would have to go in and deal with the other group, you know, and look after them. (teacher, group 3)

Incidents included pupils refusing to go in to the lessons on the grounds that they believed they were not learning anything worthwhile during the sessions. Staffing changes then further complicated the situation. The teacher who had received UKRP training was made head of English – the subject in which the school needed to make greatest improvements in attainment – and a decision was taken to suspend delivery of the UKRP so that the teacher could spend more time teaching English. In addition, the headteacher, who had been very supportive of the UKRP, then left the school to become head at a newly launched Academy, subsequently also recruiting the senior manager who had been responsible for the UKRP to the new Academy. Perhaps not surprisingly, when a replacement headteacher was appointed six months later, a decision was taken not to resume the programme the following year.

While this school was the only school at which the UKRP had been formally discontinued at the time of the second case study visit and is therefore very much the exception, the case does nevertheless usefully draw attention to some wider issues. First, the programme had been organised at the school in such a way that it was highly dependent on two individuals – the senior manager responsible and the one teacher who had received UKRP training – making it particularly fragile. Second, it illustrates how initiatives such as the UKRP may be swept aside by the need to address higher priorities such as low attainment levels.

Finding a place in the curriculum

In the third year of delivery, schools had found a variety of means of fitting the UKRP into their curricula. The original UKRP programme was set out in twelve sessions lasting ninety minutes. Clearly this was not practical for most schools and so the course content was repackaged into eighteen sessions lasting one hour. This presents a challenge to schools as 18 sessions, if followed weekly, represents roughly half of an academic year. A further constraint arises from the recommendation from the course developers that it be delivered to groups of no more than 15 as the research evidence identifying a positive impact from the programme related to groups of this size or smaller.

Two obvious means of accommodating an 18 week programme were available to schools. The first option was to offer the 18 week course in fortnightly sessions spread across the whole academic year. The second was to find a programme with which to link the UKRP so that together they could be timetabled across the whole year.

The former option was adopted by three of the schools (a group 1, group 2 and group 3 school). This very simple solution clearly offered several advantages. First, it meant the UKRP could appear as a named subject on pupils' timetables. This gave the programme an identity and, having been established in the timetable, offered a degree of independence as its continuation was not dependent on any other subject. In a group 1 school following this strategy incorporation into PSHE was considered but it was felt that pupils were disillusioned with the existing PSHE curriculum and consequently providing a separate slot for the UKRP meant it would not be tarnished by association with an unpopular course. The clear disadvantage of this solution was, however, that it meant that pupils only received UKRP sessions fortnightly. A two week gap between sessions was felt by many staff to hinder the development of good bond within the group (and especially with the adult leading the group) and also that pupils could struggle to recall what they had learned in the previous session. Such problems were of course magnified when illness, holidays, inset days or days off timetable meant that groups did not meet for four or more weeks.

The strategy of linking the UKRP with another programme, was adopted by the other four schools that were continuing to deliver the programme three years after its launch. As was reported earlier, one of the group 1 schools delivered the UKRP during English lessons. While this may not appear to be an obvious home for the UKRP, the head of English remarked:

> Having been on the training myself, I feel very much that it actually fits in with English skills superbly… I think that their speaking and listening skills in English will improve as a result of the skills that they learn on the Resilience Project. (senior manager, group 1)

In the other three schools the UKRP had been linked with either PSHE or Personal Learning and Thinking Skills. In one of the group 2 schools the UKRP had been linked with a PSHE/citizenship course. Under this arrangement, pupils first completed a transition unit at the start of Year 7 focusing on a book they had read initially in their primary schools. Half of the year group then embarked on the UKRP programme. Class groups were split into three for this purpose (with a UKRP group size of about nine pupils), an advantage that could be achieved relatively easily because of the use of non-teachers to deliver the programme. After completing the programme during the spring term, the other half of the year group then received the UKRP in the second half of the school year. Thus half the year group followed the UKRP during the first half of the year, while the other half of the year group followed a PSHE curriculum, and during the second half of the year the groups swapped over.

Arguably, one drawback of this approach resulted from the introduction, in September 2008, of the new Key Stage 3 National Curriculum and in particular its seven cross-curricular dimensions. The senior manager interviewed described the Key Stage 3 curriculum as:

> Divided up into half-termly themes… in English and Humanities, and each half-term has a PLTS which underpins that theme… an example when the kids first get here in the first year is something like 'let's stick together', so that involves introducing themselves… and making friends and all those… social skills… And then in English they do an autobiography… and humanities in the first term is all about [the local

area] – self identity and local identity [is the cross-curricular theme]. (senior manager, group 2)

As a consequence of the different cross-curricular theme for each half-term, the theme-based content of the PSHE curriculum followed by Year 7 pupils would then differ according to whether pupils followed the UKRP in the first half or second half of the year. While there may be not be good reasons to want or expect all pupils to follow an identical curriculum, it does appear that the operation of the UKRP at this school may illustrate the difficulty of incorporating atypical programmes into theme-based curricula.

In another group 2 school a slightly different solution had been found. In response to the new Key Stage 3 curriculum, the school had introduced a separate lesson to teach PLTS as a discrete subject on the school timetable. Under this arrangement, in both Year 7 and Year 8 pupils received one lesson each week in PLTS. The UKRP curriculum was followed within PLTS sessions for one term during Year 7, and this was seen as addressing some of the emotional and cognitive elements of PLTS. The other two terms were then devoted to, for example, identifying and developing learning styles. The UKRP element was delivered on a rota and thus pairs of UKRP teachers (delivering the UKRP to two halves of a class) would deliver the programme to three groups over the course of the year. In Year 8 the UKRP element was replaced by units taken from the SEAL teaching materials. One consequence of this arrangement was however that the school had slightly condensed the UKRP, by removing what staff saw as some of the repetition within the programme, so that it could be delivered in between ten and twelve one hour sessions.

The final school that had integrated the UKRP into a PSHE programme was a group 1 school and had undertaken this process in a particularly thorough manner. The process had been driven forward by a new, and very committed, UKRP coordinator. This teacher, who was also head of Year 7, had initially been given the task of organising the delivery of some SEAL materials during afternoon registration. However, the teacher had immediately seen the overlap between SEAL and the UKRP and the inadequacy of the time allocation, and so had persuaded the headteacher that one lesson per week should be allocated to a combined UKRP/SEAL programme. This programme was then delivered by a team of teachers to whole class groups (of up to 22 pupils) rather than the group size of up to 15 pupils recommended by the UKRP course developers.

The teacher explained, the announcement that PSHE was to become a statutory element of the curriculum had been a ‘godsend’ – for example in persuading the headteacher of the value of the new programme - as UKRP and SEAL could together address most of the PSHE curriculum. The examination of the two curricula, for example to identify any conflicting messages, had been undertaken in collaboration with one of the lead UKRP trainers and with advisers from the local authority.

Interestingly, during the 2007 case study visits there had been an assumption within several schools that UKRP and SEAL would be competing for the same curriculum space. However, it could be argued that in these two schools the two programmes had been mutually beneficial. In both schools it was stated that in combination (in one case also in combination with PLTS) the two had produced a programme of sufficient magnitude to command a place in the school timetable.

No attempt was made to explicitly link the UKRP with SEAL in any of the other case study schools although there was some variation in how SEAL had been implemented across these schools. The school with the most established SEAL programme, which was also the case in 2007, operated with SEAL objectives (now integrated with PLTS objectives) for every lesson within the school, organised according to half-termly themes. Meanwhile in two of the other schools SEAL teaching materials were used in PSHE-type programmes.

Interestingly, two PSHE coordinators, responsible for SEAL and who had also undertaken the UKRP training, were interviewed. They were therefore particularly well-placed to make comparisons between the two programmes. The views they expressed were nevertheless echoed by several other interviewees. For example:

> I really believe that SEAL begins to scratch the surface, and resilience is a much deeper more meaningful skill that they need to use (PSHE coordinator, group 1)

> I do think that the skills in the UKRP are more my cup of tea [than SEAL] because they're very explicit skills… The learning on the UKRP is much more overt; it's more hidden in the SEAL. The children aren't quite sure sometimes what they [are learning]. (PSHE coordinator, group 2)

However, while it was quite common for interviewees to favour the depth of learning and explicit skills of the UKRP, in other respects the SEAL programme was favoured over the UKRP by the second interviewee quoted above:

> I quite like the SEAL stuff and I think the lessons are a lot more active. I mean the UKRP lessons can sometimes be very bland and very boring and we'll desperately try to think of ways to make them more exciting. You know, there's a lot of – because there's a lot of non-activity stuff and when there are activities, the activities are the same as the activity they've done before or very similar, so I don't necessarily think that the methods that the children learn in UKRP are the best. Whereas in SEAL there seems to be a range of activities for them to do, you know, there's the – the teachers are enjoying teaching the SEAL stuff. Yeah, it's just there's a wider range of activities to do. (PSHE coordinator, group 2)

At three of the schools however SEAL had seemingly not made much of a lasting impression. For example, in describing in some detail the school's PSHE-type programme (a programme for which the manager was responsible) one senior manager stated, 'I would hope that elements of the SEAL programme are in there' implicitly acknowledging that no attempt had been made to ensure that they were part of the programme. At another school an interviewee commented somewhat bluntly:

> It just seems to have died a death… [in the] last two years you heard 'secondary SEAL', 'secondary SEAL', 'secondary SEAL'. Now I don't know when I last really heard it mentioned. And I don't know if ever very much particularly went on [at this school]. (teacher, group 2)

Interviewees' views on the UKRP

Most interviewees spoke positively about the UKRP, and this was particularly so among those working in group 1 schools. For example, when asked about any impediments to the sustainability of the programme at the school, an interviewee replied:

> It's going from strength to strength, actually. Definitely. (non-teacher, group 1)

A number of informants spoke of the improvements that had been made in their delivery of the UKRP and also its organisation over the three years in which it had been running. For example as an interviewee from another group 1 school said:

> *Would you say the sessions have improved or deteriorated or stayed the same over the last 3 years?*
> I'd say improved considerably… I think particularly [those of us] who did the original training are far more confident. We're far more confident with all of the materials…
> *And so you're more confident… do you see any difference in the pupils? Or in the groups?*
> …we can engage the kids a lot more and I think they're more comfortable with it and I think we're more likely to sort of do things spontaneously, whereas in the first year it had to be the way the manual was written. (non-teacher, group 1)

An interviewee from a third group 1 school explained that they hoped to see the programme expand at the school in future:

> I think the way we've got it at the moment is very good, but I would like to have… booster sessions… so for example this Year 7, when they go into Year 8, they [would] do three or four booster sessions, and the same with nine, ten and eleven, so you're constantly building on those skills and they're not being allowed to forget them. (teacher, group 1)

Delivering the programme was reported to be popular among those interviewed, and in only one school was reference made to a facilitator who delivered the programme but would prefer not to do so. It should also be noted that not all UKRP-trained facilitators were delivering the UKRP in 2009-10. In some cases this may have been a matter of preference although this was most often attributed to the demands of the timetable. At one school, a programme manager did however state that some UKRP-trained facilitators would not be allowed to deliver the UKRP at the school because they were thought to lack the necessary empathy and understanding of the course.

Interviewees also reported the popularity of the programme among pupils. For example:

> *How did you feel after you'd done the training? Coming back, and the prospect of running groups for the first time?*
> Worried at first, obviously, because I'd never taught, and obviously I'm learning as I go because I'm not teacher trained. But I'm just… overwhelmed at how they responded, and really enjoy it. And I know they do as well because Year 7s, they have interviews to see how they're getting on and they even say to teachers, when they get asked what's their favourite lesson, and lots of children have said resilience. (non-teacher, group 3)

Similarly another interviewee remarked on the popularity of the course among pupils:

> [Something] that came out of the questionnaires [administered by the school] last year was that some of the kids wanted it every week rather than once a fortnight. (non-teacher, group 1)

Interviewees did not in general, however, blithely claim the programme had produced a great impact on pupils' behaviour or coping strategies. For example:

> Unfortunately I can't say I do feel like any of the children use the skills they've learned outside of the session. If they're in a session, they'll do it on the board… but when they are in an actual live situation where you're in a confrontation… like with a teacher, it kind of all goes out [the window] (non-teacher, group 2)

However, some interviewees did provide examples of instances when pupils had at least reported that they had made use of techniques they had used during sessions, or where the interviewee perceived progress within UKRP sessions:

> Last year two of the boys, who had almost got into a fight, came to find me to say 'we thought about it beforehand and we walked away because we knew what was going to happen' – and that, for me, was great because they were two boys who would get into a rough and tumble… 'Be proud of us, miss' they said. (teacher, group 1)

> A couple of times some of the girls have come up to me, you know, with a problem… and I've asked them how they've dealt with it. And they've said, 'well, I remembered such and such, Miss, so this is what I did' so, you know, it's fantastic. (non-teacher, group 1)

> Two boys were having an argument [in the session]. One said 'Well, I don't like it when you put me in the middle, between you and this other kid', and [the other boy] was like 'Yes, I know, but der der der', and [the first boy] went 'No, you're not listening to what I'm saying' – well, he would never have said that, at one point. He's saying 'I'm trying to tell you how I feel about it, it's not what I want you to do about it, or about this other person, I'm telling you how I feel.'… I don't think it's going to be life-changing for them, but it's just little incidents like that where I think 'Oh, it's been worthwhile'. So I was quite pleased with that. (non-teacher, group 2)

It is also noteworthy that most interviewees spoke very positively about the potential value of the UKRP to pupils and most stated that they themselves used the UKRP skills that they had been taught which, in itself, indicates a degree of belief in the value of the skills they were teaching.

Dealing with emotive issues in UKRP sessions

It was clear from the interviews that the character of UKRP sessions, or certainly the type of real life material discussed during UKRP sessions, could vary substantially. During sessions, pupils are invited to describe real life problems to which UKRP skills may then be

applied. This was thought to be of value by most facilitators as this exchange between two facilitators exemplifies:

> *[Facilitator 1]* I think the children listen when they're talking about their own personal problems. Like we had an example today and we got a bit, well some of the kids sort of got a bit silly and a bit giddy and then we had a student and he was talking about his own personal experience and the kids were like chatting and whatever and he was like 'look, you know, I'm basically telling you something that's quite personal here you need to listen' and they did. They stopped talking and listened to him and I was like good on him for doing that. So yeah I think they respect when they're talking about their own problems, rather than you know, the ones from the book.
> *[Facilitator 2]* I think it goes the same with us as well…[if] it's about your own personal life [they say] 'oh my gosh miss'. But when you see the text book ones it's just 'yeah' *[unexcitedly]*. They just know it's not realistic and it's quite Americanised anyway so, [the text book characters] saying 'I'm a loser' - it's not real to them… (non-teachers, group 2)

Another facilitator however reported that pupils had been reluctant to acknowledge experiencing any problems:

> [They say] 'Well, we don't have any real life problems', and I think it's just because they don't want to admit to having any problems… I get the odd, every now and again, 'Oh, my sister pinches all my things out of my room, but I always get blamed for it, blah, blah, blah' but the more deep problems they're just not willing to own up to having any problems. (non-teacher, group 2)

However, this was far from being the case for all facilitators although, where pupils did provide examples of problems from real life, inevitably the type of problems that could be raised by pupils could also vary widely:

> Basically… the group I had yesterday… the kids [were] asked to identify a problem. One girl's talking about the death of her dad a couple of months ago. Another kid was talking about [the problem of] getting tickets for a… football match. (non-teacher, group 1)

When faced with genuinely distressing problems, the interviewer was at times impressed by the ability of some facilitators to describe how they had been able to use the example of the problem to illustrate the use of a UKRP skill. For example, a facilitator described a discussion focusing on the ABC model:

> …the adversity was the dog had died… it was quite hard getting them to [identify the] adversity, belief and consequence and then –
> *What's the belief, just –*
> The belief? My dog has died, 'ah this is so sad and horrible', 'why did she have to die, it's so sad'. And then the consequence was feeling sad, upset and heartbroken. So then we sort of talked about it. I don't know, we talked about alternatives and then [the pupil's] alternative was 'she was ill and she is now out of her pain and that she would be happy in heaven'. And that made him feel a little bit relieved… And I

> just keep drumming in to [them] that what you think can always be changed and it will always have an effect on how you feel. (non-teacher, group 1)

However sometimes the problems raised by the children could be of such magnitude that the facilitator could feel unable to help. In the passage below the facilitator identified both the positive and negative feelings that arose for the facilitator as a result of running UKRP sessions. In particular, the interviewee reported the appeal of feeling close to the children that arose from having them share their problems, both during sessions and subsequently outside UKRP sessions. But the facilitator also identified feelings of powerlessness and guilt that resulted from feeling unable to help and from UKRP sessions leading one particular pupil to repeatedly focus on the death of a parent. As she described:

> It's nice that they can come to me [to talk] but obviously I haven't always got the answers that they need.
> *What kind of problems?*
> …Problems with obviously not getting on with your parents, which is quite normal in Year 7. There's one girl who thinks her dad favours her sister… and so all I can do is reassure her, and 'yes, he loves you', and it's just a difficult age, she's younger so everything's sweet for her… And then we've got some looked after children and one student's worried that she might have to go back to her real parents, and she loves her foster parents. You know, they're such big problems…
> *…When they come to you with an individual problem like that, are you able to use any of the resilience skills to suggest to them then?*
> Yeah, I do. I sort of say… 'remember when we did such and such a session' and they do… But I just think some of those problems are so big, for our students, you know – that yes, they can use these skills which obviously they've learned, but they're actually asking me to really help them – do you know what I mean? To give them an answer. Which I can't do. I can do the breaking down [the problem] and looking from a different way and everything else but –
> *Right, and does it – so how does it feel if you give them that kind of answer?*
> … I love the fact they can come to me, I love the fact that they want to speak to me, but I also find it frustrating that I can't help them… as much as I'd like… I've got one student who often cries because her dad died. And when we talk about problems [in UKRP sessions] the first thought that comes into her head is her dad dying. Well, obviously then I can talk to her and say… 'we all have to deal with death' and sort of do it that way, and… 'it's really sad' and 'he wouldn't want you to be upset'. But it's a problem that always – every session that we talk about problems. And I just feel like, you know, 'am I making this student think these depressing thoughts?' and there's nothing I can do. (non-teacher, group 3)

It might be that some problems are too personal and too distressing to be discussed during a UKRP session and schools and facilitators must make decisions, in such cases, as to whether the UKRP may be helpful and appropriate. It is clear that facilitators could sometimes be faced with having to make sometimes difficult judgements about how to respond to the problems identified by pupils during UKRP sessions.

Views on the UKRP course materials

As in 2007, and as is reflected in the earlier comments relating to the SEAL teaching materials, some interviewees expressed the view that the course materials for the UKRP could be improved. For example, one interviewee distinguished between the quality and value of the concepts being taught on the one hand and the quality or appropriateness of some of the written course materials on the other:

> The skills, and what you do on the [white]board – sort of like around group discussions, or when you're introducing the models… – they respond really well to that… the ABC model that you talk about on the board, and talking about thoughts, and feelings, and consequences… When you do it on the board as a group activity, they get it fine. But when you go to do it in their book, the pictures don't relate to the ABC model; they don't have A, B, C on the top of the boxes, for example…. It's just three boxes, so they don't make that natural link between the pictures – they don't understand what they are. There's one of them, for example, that's meant to be a coach, like an American soccer coach, that's got like a chalk board with numbers and letters on and things, and the kids see it as like a maths diagram, not as a sports tutoring device. (non-teacher, group 2)

At some schools additional resources had been used to make what staff saw as improvements to the materials:

> I think the first year… we stuck rigidly to the programme, but I think we all found that very difficult. But inhibiting as well because I think the materials we felt weren't brilliant for UK kids. Everything was Americanised… [For example also] as soon as you mention "self talk" the kids think in terms of talking to yourself, but we have another resource [not produced by the UKRP course developers] which talks about ANTs… - Automatic Negative Thoughts. So things like that run much better… we've followed the programme but maybe used other resources, other than the ones that were originally provided by [the course developers]. (non-teacher, group 1)

Schools differed in their understanding of the tight copyright arrangements relating to the UKRP and at times an uneasiness was perceptible during interviews at a minority of the schools. The course materials are protected by copyright and users are not allowed to edit them. In addition, a licence to use the materials is only granted to those who have attended training events. At one school this was presented as an obstacle to maximising the value of the programme, when asked about any links between the UKRP and SEAL the interviewee replied:

> I mean obviously we're not allowed to touch anything to do with the resilience project, that has to stay, and we – that is only taught to the children who are chosen for the resilience project, and we don't use – because we're not allowed to, because it's all copyrighted by [the developers], so we don't touch any of that material with any of the other children… (teacher, group 3)

In short, for some interviewees there appeared to be a tension between the copyright and licensing constraints relating to the programme on the one hand and their appraisal of the

quality of the course materials on the other. For example, the absence of differentiated materials for lower ability pupils was pointed out by several interviewees.

Interviewees explicitly expressed their understanding that the UKRP should be delivered in full and as laid out in the teaching manual because research evidence relating to its efficacy only applied to such a mode of delivery. For example, as one interviewee remarked:

> The programme is so highly structured and I think that worked well, and that's not something we want to mess around with… because my concern is that none of us are trained psychologists or anything like that… we're not qualified or in a position to start altering the content of something that is clearly long established and obviously works. (Teacher, group 1)

Yet on the other hand some interviewees also felt that the materials could certainly be improved upon. For example, at one school, changes had been made to the physical materials provided to pupils. For instance, attempts had been made to further 'de-Americanise' some of the resources. In addition, the UKRP files - containing for example, cartoons with thought bubbles to complete, scripts for role plays, as well as descriptions of lesson plans – had been reorganised. Material that was deemed to be unhelpful for the pupils (for example wordy lesson plans) was removed and what remained was separated into a more user-friendly homework booklet and classwork booklet.

This tension was particularly clear for schools that wanted to introduce UKRP booster sessions in Year 8 or Year 9 but did not have appropriate materials with which to do so. A local authority coordinator spoke explicitly about the licence agreement:

> we have to sign a licence agreement with [the course developers] so that… all our members of staff that do the training know that that licence is for them to teach the Resilience programme as it is to young people that they work with and that… they can't start adapting it… and we take that seriously, you know, it's their product and we absolutely respect that. (local authority co-ordinator)

The interviewee then spoke in positive terms about the good relationship with the course developers and of their understanding of the local authority's needs. It was however also evident that the responses to those needs were not always in keeping with the needs of schools. For example some schools wished to introduce UKRP booster classes for Year 9 pupils who had previously followed the UKRP programme in Year 7. However they were not satisfied with the examples of booster exercises offered by the course developers. Consequently experienced UKRP facilitators had developed their own materials, as the interviewee explained:

> [Two facilitators] are writing their own. And we will share them, so they'll run them with their groups this year. We told [the course developers] that that's what we were doing, you know, we'll share the material with them once we know, you know, whether they've been any good or not…but [the facilitators] are both people who have been involved in this for a long time, so they're both trained up to facilitator level and they [have] taught it lots and lots of times. So they've got a really good understanding about it. So you know, again I would be concerned if we had a brand new teacher that was suddenly saying, you know, ['I'm going to write some new materials for this

> purpose'] but the fact [they]… really understand it well, I'm quite confident… (local authority coordinator)

The existence of such an ad hoc arrangement (notably an open and cooperative one) suggests there may be a mismatch between the priorities of the course developers and the schools. That is, the *timely* production of new materials to reinforce learning is a priority for schools.

The tension between the prescriptive nature of the programme and conventional approaches to teaching was perhaps particularly clear when two senior managers each compared UKRP lessons with the style of lessons expected by OFSTED:

> the prescriptive nature of [the UKRP] is not particularly how teachers work nowadays. You know, the best teachers respond to what's actually happening in their lesson, rather than – you know, you'll put together a lesson plan, for example, on any given day, and let's say for example an OFSTED inspector came in, and you've got a lesson plan in front of you; they'd expect you to respond to what's actually going on in the lesson, rather than sticking neatly to your plan (senior manager, UKRP-trained, group 3)

> I just think that that because of the nature of the course, you've got to really work hard… all the Ofsted criteria is about pushing the progress agenda isn't it? … So to show progress in this sort of lesson I think is a real challenge, even the most experienced teacher… [When I observed a class] it was very, it's a very prescriptive course isn't it? And [the teacher] delivered a lesson that reflected that. You know, did it very well, did it very well and got a lot of the kids [engaged], but it wasn't Ofsted standard - to hit the higher notes. (senior manager, group 2)

Interestingly, another interviewee suggested that, given the lack of expertise in evaluating the quality of UKRP delivery, it would be valuable if the course developers could observe some UKRP sessions to provide feedback as to whether the programme was being delivered in the manner the developers had in mind.

It is important to note that the prescriptive nature of the programme was not viewed in an entirely negative light by respondents. Indeed, the second of the senior managers quoted above also remarked that the prescriptive content offered the advantage of greater confidence for senior managers in knowing what was being delivered compared with more flexible courses (and in this instance a contrast was drawn with the more flexible SEAL programme materials). In addition, as was noted earlier, some interviewees also acknowledged their lack of expertise in psychology and therefore their competence to deliver the programme in a more flexible manner.

In many respects, the interview data echoed that collected during the first round of visits to case study schools in 2007. For example, the training events for new group facilitators carried out in the summer of 2008 (held in Cambridge) and summer of 2009 (held in each of the three local authority areas) were almost universally described in very positive terms although there was perhaps less missionary zeal evident among interviewees than had been the case in the first set of interviews, after the first training event in 2007 (in Philadelphia). The training was nevertheless compared favourably with other training experienced by

interviewees. Perhaps most importantly of all, among the few interviewees who had attended more than one training event (for example because they subsequently played a role supporting the lead trainer in another training event) there was no suggestion that the quality of training had diminished over time. Indeed, more than one even suggested that it had improved although it was also observed that the shift from providing residential courses to locally-based training events meant trainees were less immersed in the training programme (for example having less time to discuss the programme outside sessions).

Conclusion

We have seen that, in its third year, the UKRP was being delivered in most of the case study schools. At seven out of the nine schools visited the programme was being provided to the whole Year 7 intake. This is undoubtedly the best indicator of schools' commitment to and satisfaction with the UKRP during the first three years of its operation. In addition, more staff had received UKRP training at each of those schools and the programme had been embedded in the school curriculum.

Schools varied as to whether the programme was delivered primarily by teachers or by school support staff and also in the proportion of trainees drawn from these groups. At several schools interviewees reported a lack of demand for the role from among the teaching staff compared with the support staff. There appeared to be several reasons why this might be the case including workload, pay arrangements and whether the UKRP was seen as good a career development opportunity.

Schools at which there was reportedly continuing demand from teachers to attend UKRP training, and in which the programme continued to be delivered primarily by teachers, had good records of pupil progress to GCSE level, and this may have enabled teachers or managers to give greater priority to the UKRP. This was not however the case in all of the more academically successful schools and clearly in one school, senior managers had not accommodated the staffing and logistical demands of the UKRP that would have been required for it to be delivered it to the whole year group. In general it appeared that management backing for the programme and strong and persuasive leadership of the UKRP within individual schools could be vitally important.

Schools had found different ways of accommodating the UKRP within the curriculum. In some schools it appeared as a named lesson on the school timetable. While this could offer an identity and a degree of stability to the UKRP within schools it also required the programme to be taught once per fortnight rather than weekly – and this is not the preferred mode of delivery.

In other schools the UKRP had been linked with other programmes of study. At one school, the UKRP was delivered by English teachers and one lesson per week was devoted to the UKRP for part of the year. At other schools the UKRP had been linked with PSHE, PLTS or SEAL in order to create a programme of sufficient size to be timetabled at least once per week across the year. At one school the UKRP had been condensed so that it could be delivered in a single term while at another it was delivered to larger teaching groups than had originally been recommended. Indeed, if we consider 18 weekly sessions delivered to

groups of no more than 15 pupils as the ideal mode of delivery for the UKRP, this mode was only achieved in two of the nine case study schools in 2009-10.

In many ways, the responses of interviewees were similar to those reported two years after the first case study visits. UKRP facilitators were generally very positive about the objectives of the UKRP and about their experience of delivering the programme. Most facilitators reported trying to use UKRP skills in real life, which in itself indicates their belief in the value of the programme. While they spoke enthusiastically about the value of the training they had received for the UKRP, some interviewees expressed reservations about the quality of the UKRP course materials.

As a whole, the UKRP case study schools show that the UKRP had been implemented in a variety of ways and that it is clear that no single model would be suitable for all schools. Many demands are made of schools and they must of course deploy their resources thoughtfully to address a range of priorities. Most of the UKRP case study schools showed continuing commitment to the programme and, in most of the schools, the UKRP has been successfully embedded within the school.

Many issues relating to the implementation of the programme were being addressed within individual schools or local authorities; however, it is perhaps useful to explicitly draw attention to five policy-related issues.

- First, the findings suggest that it is essential that the programme has the backing of senior management within schools.

- Second, a preferred model of delivery for the UKRP, based on the recommendations of the course developers and the findings of this study, might be seen as involving sessions led only by trained facilitators (and indeed this is a requirement under its licensing arrangements) to groups of no more than 15 pupils, during 18 weekly, hour long, sessions. Only two of the nine schools were delivering the UKRP in accordance with this preferred delivery model. Thus, it is clear there are pressures on the UKRP within schools, and these may arise from financial or timetabling demands or from pressure to improve pupils' levels of attainment. Tighter constraints could arise from falling school budgets or a reduction in the number of parents consenting for their children to participate in the programme and this would necessitate greater management backing for the project.

- Third, there is a drift evident in some schools towards the programme being delivered by non-teachers, in part because of the pressures noted above. Such a drift may reduce the pool from which potential trainees may be selected.

- Fourth, as we have seen, the quantitative analysis relating to the best matched experiment identified some beneficial average effects arising for a universal intervention group. However, this does not rule out the possibility

that the programme could potentially have a detrimental effect on some pupils.

- Fifth, some facilitators stated that they had enhanced some of the UKRP teaching materials and so there is the possibility that programme fidelity could be reduced if teachers do not feel the materials are of a sufficiently high quality.

5. Appendix

Questionnaires used to survey treatment and control pupils at all data collection points

Outcome: Symptoms of depression
Children's Depression Inventory
UK supplier: Harcourt Assessment
UK copyright holder: Multi-Health Systems Inc.
The version used excludes item 9, as this concerns suicidal ideation and this was not deemed appropriate or necessary.
Manual: CDI Technical Manual, Maria Kovacs 2003, pub. Multi-Health Systems

Outcome: Symptoms of anxiety
Revised Children's Manifest Anxiety Scale
UK copyright holder: Western Psychological Services
The wording of some items was modified slightly for UK English.
Manual: Revised Children's Manifest Anxiety Scale [RCMAS], Cecil R. Reynolds and Bert O. Richmond 2000, pub. Western Psychological Services

Outcome: Life satisfaction
Brief Multidimensional Students' Life Satisfaction Scale
This is not copyrighted, and is available online at:
http://www.cas.sc.edu/psyc/facdocs/hueblifesat.html
Scoring instructions are also available here.

Outcome: Behaviour
Goodman Strengths and Difficulties Questionnaire
(1) Self-report version
(2) Teacher-report version
All versions of the Goodman SDQ are available online at:
http://www.sdqinfo.com/

The Goodman SDQ can be used free of charge, although it is not possible to modify the wording. For the purposes of evaluating an intervention there is one version (an 'initial' version) that has a reference period of six months, and a follow-up version with a reference period of one month. All data collections subsequent to the baseline use the one-month reference period. When the follow-up version is used it is acceptable to change the word 'clinic' to 'classes' or 'programme', and this has been done, in order to clearly refer to the UKRP. Modifications to the format are acceptable provided the layout is essentially the same.

Scoring details and references are available at the website listed above.

6. References for Empirical Evaluations of PRP and UKRP

Brunwasser, S. M., J. E. Gillham and E. S. Kim, “A Meta-Analytic Review of the Penn Resiliency Program’s Effect on Depressive Symptoms”, *Journal of Consulting and Clinical Psychology,* 77(6), 1042-1054

Cardemil, E., Reivich, K., Beevers, C.G., Seligman, M.E.P., & James, J. (2007). The prevention of depressive symptoms in low-income, minority children: Two-year follow-up. *Behaviour Research and Therapy, 45,* 313-327.

Cardemil, E.V., Reivich, K.J., & Seligman, M.E.P. (2002). The prevention of depressive symptoms in low-income minority middle school students. *Prevention & Treatment, 5,* np.

Challen, A. R., S. J. Machin, et al. (2009). UK Resilience Programme Evaluation: Interim Report, Department for Children, Schools and Families.

http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RR094%20(1).pdf

Chaplin, T.M., Gillham, J.E., Reivich, K., Elkon, A.G.L., Samuels, B., Freres, D.R., Winder, B., & Seligman, M.E.P. (2006). Depression prevention for early adolescent girls: A pilot study of all-girls verses co-ed groups. *Journal of Early Adolescence*, *26,* 110-126.

Cutuli, J.J. (2004). Preventing externalizing symptoms and related features in adolescence. Unpublished honors thesis, University of Pennsylvania, Philadelphia, PA.

Cutuli, J.J., Chaplin, T.M., Gillham, J.E., Reivich, K., & Seligman, M.E.P. (2007). Preventing externalizing symptoms and related features in adolescents. Manuscript in preparation.

Cutuli, J.J., Chaplin, T.M., Gillham, J.E., Reivich, K.J., & Seligman, M.E.P. (2006). Preventing co-occurring depression symptoms in adolescents with conduct problems: The Penn Resiliency Program. *New York Academy of Sciences, 1094* , 282-286.

Gillham, J.E. (1994). Preventing depressive symptoms in school children. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Gillham, J.E., Hamilton, J., Freres, D.R., Patton, K., & Gallop, R. (2006). Preventing depression among early adolescents in the primary care setting: A randomized controlled study of the Penn Resiliency Program. *Journal of Abnormal Child Psychology, 34*, 203-219.

Gillham, J.E., & Reivich, K.J. (1999). Prevention of depressive symptoms in school children: A research update. *Psychological Science, 10,* 461-462.

Gillham, J.E., Reivich, K.J., Freres, D.R., Chaplin, T.M., Shatté, A.J., Samuels, B., Elkon, A.G.L., Litzinger, S., Lascher, M., Gallop, R., & Seligman, M.E.P. (2007). School-based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting and Clinical Psychology, 75,* 9-19.

Gillham, J.E., Reivich, K.J., Freres, D.R., Lascher, M., Litzinger, S., Shatté, A., & Seligman, M.E.P. (2006). School-based prevention of depression and anxiety symptoms in early adolescence: A pilot of a parent intervention component. *School Psychology Quarterly, 21,* 323-348.

Gillham, J.E., Reivich, K.J., Jaycox, L.H., & Seligman, M.E.P. (1995). Prevention of depressive symptoms in schoolchildren: Two-year follow- up. *Psychological Science, 6*, 343-351.

Jaycox, L.H., Reivich, K.J., Gillham, J., & Seligman, M.E.P. (1994). Prevention of depressive symptoms in school children. *Behaviour Research & Therapy, 32*, 801-816.

Miller, J.B. (1999). The effect of a cognitive-behavioral group intervention on depressive symptoms in an incarcerated adolescent delinquent population (juvenile delinquents). Unpublished doctoral dissertation, Wright Institute Graduate School of Psychology, Berkeley.

Pattison, C., & Lynd-Stevenson, R.M. (2001). The prevention of depressive symptoms in children: The immediate and long-term outcomes of a school based program. *Behaviour Change, 18*, 92-102.

Quayle, D., Dziurawiec, S., Roberts, C., Kane, R., & Ebsworthy, G. (2001). The effect of an optimism and lifeskills program on depressive symptoms in preadolescence. *Behaviour Change, 18,* 194-203.

Reivich, K.J. (1996). *The prevention of depressive symptoms in adolescents.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Roberts, C., Kane, R., Bishop, B., Matthews, H. & Thompson, H. (2004). The prevention of depressive symptoms in rural children: A follow-up study. *International Journal of Mental Health Promotion, 6,* 4-16.

Roberts, C., Kane, R., Thomson, H., Bishop, B., & Hart, B. (2003). The prevention of depressive symptoms in rural school children: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 71,* 622-628.

Shatté, A.J. (1997). Prevention of depressive symptoms in adolescents: Issues of dissemination and mechanisms of change. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Yu, D.L., & Seligman, M.E.P. (2002). Preventing depressive symptoms in Chinese children. *Prevention & Treatment, 5,* np.

Zubernis, L.S., Cassidy, K.W., Gillham, J.E., Reivich, K.J., & Jaycox, L.H. (1999). Prevention of depressive symptoms in preadolescent children of divorce. *Journal of Divorce and Remarriage, 30,* 11-36.

Ref: DFE-RR006

ISBN: 978-1-84775-760-9

© Enterprise LSE Ltd

June 2010