International Journal of Offender Therapy and Comparative Criminology

http://ijo.sagepub.com/

Addressing Criminality in Childhood: Is Responsivity the Central Issue?

Claire Nee, Tom Ellis, Paul Morris and Amy Wilson

Int J Offender Ther Comp Criminol published online 15 October 2012
DOI: 10.1177/0306624X12460377

The online version of this article can be found at:
http://ijo.sagepub.com/content/early/2012/10/11/0306624X12460377

Published by:
SAGE
http://www.sagepublications.com

Addressing Criminality in Childhood: Is Responsivity the Central Issue?

International Journal of
Offender Therapy and
Comparative Criminology
XX(X) 1–27

DOI: 10.1177/0306624X12460377
http://ijo.sagepub.com

Claire Nee[1], Tom Ellis[1], Paul Morris[1], and Amy Wilson[1]

Abstract

The responsivity principle is the third element of the now well-established risk–need–responsivity (RNR) model of offender rehabilitation. Accruing evidence suggests it is often sacrificed in intervention programs. We aim to demonstrate the central importance of this principle when designing offender interventions by describing the results of a successful, highly responsive intervention for very young children (aged 7 upward) who have offended. A small slice of the offending population as a whole, child offenders are nevertheless tomorrow's serious, violent, and prolific lawbreakers, yet little is understood about what reduces their risk. Recent developments on responsivity are reviewed, before presenting the evaluation indicating significant and sustained drops in risk of recidivism. In-program factors such as the nature and dosage of interventions are examined, alongside outcome data. The article discusses how RNR and other models might apply to this particularly young and underresearched age group.

Keywords

responsivity, RNR model, Good Lives Model, child offenders, LSI-R, offender programs

Introduction

A number of meta-analyses have been undertaken in the last two decades on the outcomes of offending behaviour programs. They have involved hundreds of interventions for both adult and young adult offenders (for reviews, see Garrido &

[1]University of Portsmouth, UK

Corresponding Author:
Claire Nee, University of Portsmouth, King Henry Building, King Henry 1st Street, Portsmouth PO1 2DY, UK
Email: claire.nee@port.ac.uk

Morales, 2007; Lipsey, Landenberger, & Wilson, 2007; or McGuire, 2008). They have indicated beyond doubt that well-delivered programs adhering to the risk–need–responsivity (RNR) principles (Andrews & Bonta, 2003) are effective in reducing reoffending in both these age groups.

The most common program that adheres to the RNR framework for adult offenders in the English-speaking world is Reasoning and Rehabilitation (Ross, Fabiano, & Ewles, 1988) and its later derivatives such as Enhanced Thinking Skills (ETS). However, more recent evaluations of these programs and their impact on the recidivism rates when applied to large samples of prisoners, particularly in the United Kingdom, have been equivocal at best (Cann, 2006; Falshaw, Friendship, Travers, & Nugent, 2003). Many reasons have been given for this, including inaccurate matching of risk to intervention intensity (Palmer et al., 2008), lack of program integrity and program drift due to the considerable expansion of programs in the late 1990s (Falshaw et al., 2003), and inadequate assessment of readiness to change (Hollin & Palmer, 2006). Further factors are highlighted in studies of program attrition, including responsivity issues such as lack of direct relevance to offenders' lives (Cann, 2006; McMurran & McCulloch, 2007; Veira, Skilling, & Peterson-Badali, 2009), lack of sensitivity to ethnicity issues (Wormith & Olver, 2002), lack of individual tailoring of programming to offender's personal needs (Veira et al., 2009), varying intellectual abilities in groups (McMurran & McCulloch, 2007), and variation in cognitive deficits and their link to offending behaviour (Cann, 2006). Andrews and Bonta (2010) have shown that the most successful programs strongly adhere to all three RNR principles, but the latter findings suggest that responsivity is an aspect very likely to be sacrificed. From a U.K. point of view, a strong and understandable drive for program integrity through the correctional accreditation process may have unwittingly or otherwise (see Ellis & Winstone, 2002) produced a "one-size-fits-all" approach that has forfeited the contribution and importance of responsivity. The aim of the present article is twofold. First, given the research findings above, we aim to demonstrate the importance of keeping the responsivity principle firmly at the centre of one's aims when designing offender interventions. We will illustrate the potential this might have with reference to the results of a highly responsive program for very young offenders. Our second aim is to contribute to the significant gap in knowledge about effective interventions with preteenage and early adolescent offenders, with a specific focus on the "in-program" factors that Wormith and colleagues (2007) suggested necessary to really understand not only whether the intervention works but also how and why.

To realise these aims, the article first outlines the key elements of responsivity and then examines what is known about child offenders. Next, it presents data from two studies. To provide context, we first present an ongoing risk/need and recidivism summative evaluation to illustrate the effectiveness of the program. Within this context, we then present an analysis of the nature and intensity of the intervention to illustrate the levels of responsivity in the program. We aim to show how the program not only satisfies many of the general principles of "what works" for adult offenders but also raises issues that require further examination in the very young population of offenders we report on.

The Concept of Responsivity

Alongside the risk principle and the need principle, responsivity takes account of the offender's individual characteristics, such as ability to learn, motivation, and personality traits, when designing treatment programs to maximise the chances of changing for the better (Bonta & Andrews, 2007). Based on their extensive empirical and theoretical development of effective offender rehabilitation, Andrews and Bonta refined the responsivity principle into general and specific responsivity (Bonta, 1995, 2002; Bonta & Andrews, 2007). The former refers to the value and efficacy of incorporating cognitive-behavioural, psychosocial techniques into rehabilitation programs and to the general quality of the therapeutic relationship. Specific responsivity highlights the importance of matching, as far as possible, the mode and style of delivery to the characteristics of the treatment population, that is, cognitive (e.g., intelligence and problem-solving skills), personality (e.g., extraversion, impulsivity, clinical diagnoses, and mental health), and sociocultural (e.g., age, ethnicity, gender, and level of education). This approach is also consistent with evidence from psychological therapy, in that treatment effectiveness is enhanced by tailoring interventions to the individual characteristics and needs of the client (Kazdin, 2008).

Perhaps it is not surprising that responsivity is the last to be mentioned of the three principles of RNR and the least researched. It is very likely the hardest principle to uphold. Although we are now in the fourth generation of sophisticated risk and need assessment tools (Andrews, Bonta, & Wormith, 2006), ploughing down to the array of individual client characteristics and responding to these effectively (for instance, by providing more one-to-one sessions, or by implementing special and/or time-limited treatment for subsections of the population) are both complex and potentially more resource intensive/expensive in comparison with responding to risk and need alone. Moreover, as the offender changes, interventions will have to change and in turn, measures of responsivity will have to be developed to match this for effective evaluation.

Some change and improvements are taking place in the United Kingdom to address responsivity issues in programs for adult offenders. ETS has been replaced by a more strengths-based, goal-oriented program with several one-to-one sessions to address participants' individual needs.[1] Alongside this, there are new programs for female offenders and adaptations to existing programs for minority ethnic groups (Stephens, Coombs, & Debidin, 2004). These changes have resulted largely from the influence of the Good Lives Model (GLM; Langlands, Ward, & Gilchrist, 2009; Ward & Stewart, 2003; Ward, Yates, & Willis, 2012) over the last number of years. Based on positive psychology, the GLM grew out of a recognition that focusing on deficits alone was unlikely to motivate the offender to change. Although acknowledging the importance of accurately assessing risk, it aims to build on the RNR approach by tapping into the basic human goal of a good life, and by highlighting the strengths an individual has for achieving that goal.

Despite these laudable developments in interventions for adult offenders, as noted above, there has been little to fill the gap in knowledge about what might be effective

with child offenders (Garrido & Morales, 2007; Loeber & Farrington, 2001). This is surprising, given the evidence that effect sizes for interventions are stronger for high-risk teenagers than for adult offenders (Redondo, Sánchez-Meca, & Garrido, 2002). The extent to which the impact might be even greater for younger children therefore is clearly a key area of investigation. We do have a considerable body of work on *why* children offend (see Farrington & Welsh, 2007; Stouthamer-Loeber & Loeber, 2002) but most of the studies carried out on effective interventions relate only to the narrower clinical field of problematic children and adolescents with conduct disorders (see Loeber, Burke, Lahey, Winters, & Zera, 2000; Lynam et al., 2009; Singleton, Meltzer, Gatward, Coid, & Deasy, 1998). There is also some evidence on effective ways of preventing criminality in young children who might be "at risk" of doing so,[2] but who are not yet officially deemed to be offenders (see Farrington & Welsh, 2007, for a review). There is clearly a need, therefore, to take a broader view on children who offend.

Although preteenage child offenders are a small slice of the offending population as a whole, there is strong evidence to show that among them are tomorrow's most serious, persistent, and violent offenders committing disproportionate amounts of serious crime when compared with those who begin offending in adolescence or later (Moffitt, 1993; Odgers et al., 2007). Only two projects in recent years, one in North America and one in the United Kingdom, have produced evaluations that directly measure risk of reoffending and delinquency in children already known to be persistently offending (Koegl, Farrington, Augimeri, & Day, 2008; Nee & Ellis, 2005). Furthermore, only one of these evaluations (Nee & Ellis, 2005) incorporates a validated risk assessment tool (the Level of Service Inventory–Revised [LSI-R]; Andrews & Bonta, 1995).

The evaluation reported here builds on the earlier outcome evaluation from Nee and Ellis (2005). Many researchers in this field have called for more research on the extent to which interventions are responsive to the individual characteristics of the participants (Andrews & Bonta, 2010; Hollin, 2006; McGuire, 2006; Wormith et al., 2007). This is known to have a strong impact on the likelihood of success (Andrews & Bonta, 2010). In the current evaluation, therefore, we have incorporated both outcome measures and examined the nature of the program to understand how it might work.

Based on Nee and Ellis (2005), we predicted significant reductions in risk in the outcome evaluation (Study 1). Study 2 looks more closely at the in-program responsivity factors. This was an exploratory study with this age group, though we did expect that reductions in risk would be associated with increased levels of input and matching of input to participants' individual needs (Koegl et al., 2008; Veira et al., 2009).

Study 1—Outcomes of a Continuing Risk/Need and Recidivism Evaluation

Method

Design. The study had a mixed factorial design involving both repeated measures and comparison with an untreated group.

The Program. Referrals to the program come from social workers, the local Youth Offending Team (YOT), the Education Department, a local community safety partnership, parents, and from participating offenders themselves. The criteria for referral and acceptance onto Preventing Youth Offending Program (PYOP; in line with Youth Justice Board recommendations) are as follows:

- Prolific offenders, defined through the national Youth Offending Information System (YOIS) database, as anyone with 10 offences in 12 months or anyone facing a custodial sentence;
- Offenders with special needs, such as sex offenders; and
- Preventative/protective referrals for young people aged between 7 and 12 (while participants of this age were too young to be officially involved with the criminal justice system, all were known to be offending by the local police and the project coordinator).

A community-based (nonresidential) program, PYOP intervention includes one-to-one mentoring for reintegration into education, anger management, and constructive use of time. There is group work for antisocial behaviour, problem solving, anger management, victim awareness, interpersonal skills, substance misuse, appropriate sexual behaviour, and health issues. There are also art, music, and drama workshops plus challenging outdoor activities to develop self-esteem and social skills. These are done on either a one-to-one basis or in groups dependent on the needs of the child and are mostly delivered by trained project workers with backgrounds in psychology or social work, with some volunteer input in the group-work sessions. Siblings are welcome at most of these provisions, and counselling and referral are available to parents. Group work is based at the organisations' headquarters, but one-to-one work can occur in any setting, including the child's home. Key elements of successful interventions for teenage offenders, noted by Lipsey and Wilson (1998), such as interpersonal skills, individual counselling, and multimodal and cognitive-behavioural elements are all key components of PYOP and it fits well with the positive, goal-oriented approach of the GLM (Ward & Stewart, 2003). Where needs are identified that cannot be addressed by core resources, further expertise is bought in. No two programs are identical, with a mixture of one-to-one sessions and group work usually on a weekly basis, indicating high responsivity, and each participant's case is reviewed every 8 weeks, with input and dosage adjusted as necessary. Participants are continually monitored by the project coordinator who receives at least weekly feedback from those observing the participants, including police officers, parents, schools, social workers, and program workers.

Participants. Since its inception, PYOP has worked intensively with young people known to be persistently offending. At the time of the current evaluation, the intervention group consisted of 67, with a comparison group of 24. The average age for the group was 13 (SD = 2.58), with a range from 7 to 17. Participants were characterised by a wide range of risk factors typical of this group (Farrington, 2002). All but 4 were

male and overwhelmingly of White British ethnicity. Ninety percent had been excluded from school either permanently or temporarily and there were high levels of local authority care, drug and alcohol misuse, and inclusion on the Child Protection Register. Eight had already served custodial sentences.

Sixty percent had property crime as a current offence, whereas nearly a third were associated with more expressive offences, such as criminal damage and assault, including two very young offenders allegedly involved in armed robbery and arson, respectively. Of the 67 participants evaluated, 7 had allegedly committed sex offences, including rape and indecent assault. We present analysis on the total pool of the intervention group (Group 1, $n = 67$) and two further subsets of this group. Group 1a comprises 46 participants who had received treatment for a year or more and Group 1b comprises 24 participants who had stayed with the project for 2 years or more. One of PYOP's unusual features is its long-term approach to intervention, and these analyses demonstrate both the impact and challenges this type of approach can involve.

Comparison Group. The focus of our research, namely, preteenage and early adolescent offenders, involves a rare and difficult-to-reach group, which is consequently underresearched. Thus, it was highly unlikely that we could identify and engage a comparison group from the general or offending population. Indeed, the difficulties of achieving a randomised control group in forensic settings has been described in detail elsewhere (Farrington & Jolliffe, 2002; Hollin, 2006, 2008). Consequently, our research design adopted the "incidental" comparison group approach recommended by the above authors and also used in analogous studies of another difficult-to-reach group, that is, sex offenders (Marshall & McGuire, 2003). As a minimum standard, this design requires that the comparison group constituents are not treated and are approximately matched on key evaluation features. This group of 24 was originally assessed and accepted onto the program, but did not engage and had little or no intervention. We deemed these individuals as essentially untreated. It could be possible that this comparison group was different or more challenging from the intervention group because they did not engage on factors such as risk, seriousness of offending behaviour, or other demographic features. However, we found that they were highly comparable with those who continued with the project in terms of age, M age = 14, SD = 1.39, range = 11-16, $\chi^2(2) = 0.423$, *ns*; gender (all were male); educational background (all but 1 had been excluded from mainstream school, either temporarily or permanently); criminal behaviour ($n = 10$ burglary, $n = 10$ theft and criminal damage, $n = 4$ for violent offences, 2 had already served custodial sentences); and most importantly risk. There was no significant difference between their risk/need scores as measured by a robust assessment tool involving 10 static and dynamic factors (see below) at the time of embarking on the project (intervention group: $M = 19.67$, $SD = 8.5$; comparison group: $M = 19$, $SD = 7.9$; $p = .735$, *ns*). We therefore considered that they reached Level 4 of the Scientific Methods Scale as they approximately matched the intervention group on key evaluation features (Sherman et al., 1997). Given the accruing evidence on readiness for change in offenders (Burrowes & Needs, 2009; McMurran,

2009; Ward, Day, Howells, & Birgden, 2004), it is likely that the lack of engagement in the comparison group could be explained better by the fact that they were not independently ready to engage at that time. Comparison group participants intermittently attended the YOT service that was located in the same building as PYOP and were approached on an ad hoc basis by the project coordinator around 6 months post their first LSI-R to complete another LSI-R assessment.

Assessment Tools

LSI-R risk assessment tool. The robustness of the LSI-R in predicting risk of reoffending in the male adult offenders for which it was designed has been well documented (Andrews & Bonta, 2003; Gendreau, Little, & Goggin, 1996; Raynor, Kynch, Roberts, & Merrington, 2000). With regard to its psychometric properties, numerous studies have indicated both high interrater reliability and test–retest reliability (ranging from $r = .80$ to .99; Hollin, Palmer, & Clark, 2003). Similarly, its predictive validity with regard to reconviction has been repeatedly shown to be very high. For instance, in a large U.K. sample of offenders, LSI-R assessments correctly predicted 65.4% of cases compared with 67.1% correct prediction by the Home Office Offender Group Reconviction Scale (Version 2; Home Office, 1996), which uses static predictor variables only (Raynor et al., 2000). In the last decade, it has grown in popularity for use with a wider variety of offenders and subgroups thereof including female offenders, violent offenders, and sex offenders (Hsu, Caputi, & Byrne, 2009). A youth version of the LSI-R has been published since the onset of our evaluation, but has not yet undergone as thorough an analysis of its ability to predict risk as its counterpart for adults. Incorporating a strong risk of reoffending predictor was of central importance to our evaluation given that a large proportion of our sample was below the age of criminal conviction and was therefore ineligible for a formal reconviction study using official records. For this reason, and for consistency purposes, we have continued to use the LSI-R that appeared to work well with younger populations (Nee & Ellis, 2005) and has since been used successfully for risk prediction in conduct disordered adolescents as well (Butler, Fearon, Atkinson, & Parker, 2007). Unlike many other risk predictors, it is based on a strong (empirically based) theoretical model of criminality (Hollin, 2002) and with 10 subcomponents reflecting both static and dynamic contributory factors, making it a versatile tool for both researchers and practitioners.

YOIS. As a background behavioural measure of the effects of PYOP on reducing reoffending, for both intervention group and comparison group, we had access to YOIS that records all charges made by the police against children over the age of criminal responsibility (10 years) in the United Kingdom. It is worth pointing out that this data is a less reliable measure of the impact of PYOP than the LSI-R data as it is, paradoxically, both incomplete and overinclusive (the majority of charges are very minor and do not result in conviction) in comparison with other official offending records. It is therefore only useful as an additional, confirmatory (or otherwise) measure of the impact of the intervention.

Procedure. An LSI-R assessment (requiring both interview and examination of historical files to produce a risk score) was carried out on each new participant as he or she entered the intervention (Time 1) and then repeated every 6 months to assess change. Subcomponents within LSI-R were also analysed to see exactly where contributions to change were strongest. Those in the comparison group were known to the project coordinators and intermittently attended the YOT's offices in the same building as PYOP, which enabled a Time 2 LSI-R assessment (6 months post Time 1) with participant consent. This allowed a comparison between this group and the intervention group for this time period. These assessments were primarily used as an evaluation tool, but were also informative for the practitioners working with the participants in terms of responding to risk and need alongside the use of other monitoring tools.

The YOIS database was examined for all intervention and comparison group participants and the number of charges for the 6 months before PYOP intervention was recorded, along with the first 6 months of intervention, the second 6 months of intervention, and so on, so that police charges data mapped directly on to the 6-monthly LSI-R data collection periods. At the time of initial assessment, written informed consent was acquired from parents and participants allowing the use of participants' LSI-R data, their YOIS printouts, and access to their personal files for the purposes of the evaluation.

Results

We first present an analysis of the changes in risk/need scores of the full cohort of participants ($n = 67$, Group 1) in comparison with the untreated group ($n = 24$) between the start of the program (Time 1) and 6 months later (Time 2). As the program is unusually long term in its approach, we are then able to present a within-participants analysis for those remaining in the treated group after a year of intervention (Group 1a, $n = 46$), and those still involved after 2 years of intervention (Group 1b, $n = 24$), to assess the impact of the intervention over a longer time period. Between groups, comparisons were not possible for the latter two analyses as it was not possible to trace the comparison group after 6 months to take assessment measures.

We then present a similar set of analyses of changes in police charges data as a proxy measure of the impact of the intervention in reducing reoffending, that is, a comparison of the first 6 months of charges between the treated group and untreated group and then two further within-group analyses of the treated group only, after a year and 2 years of intervention.

Risk/Need Analysis. The primary aim of the analysis was to assess changes within the LSI-R scores of the PYOP participants (Group 1) over time. LSI-R scores could also be compared between Times 1 and 2 with those of the comparison group. A 2 (time) × 2 (group) ANCOVA was undertaken with age as the covariate. As the age range was wide (7-17 years), we wanted to control for the effect of age. As predicted, we found that PYOP was causing a statistically significant reduction in its participants' risk of reoffending. Figure 1 illustrates a significant time by group interaction, $F(1, 88) = 20.59$, $p < .001$, $\eta^2 = .19$.

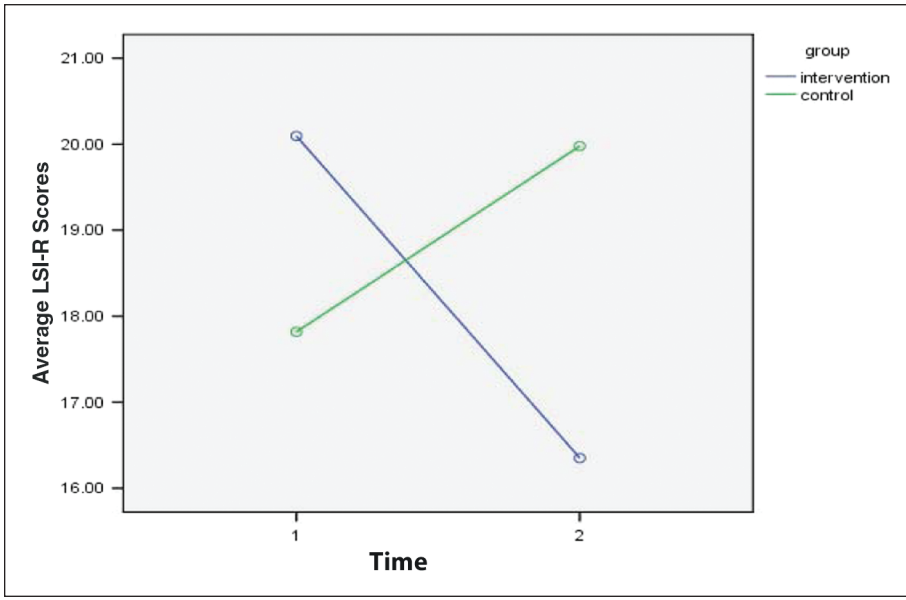

Figure 1. Average scores on LSI-R at Time 1 (initial assessment) and Time 2 (6 months later) for intervention group and comparison group (controlling for age)
Note: LSI-R = Level of Service Inventory–Revised.

Post hoc comparisons showed no significant difference between the groups at Time 1 ($t = 0.339$, $p = .735$), but a significant reduction for the PYOP group at Time 2 ($t = -2.40$, $p = .036$). After 6 months of PYOP intervention, the average LSI-R score for participants had reduced from 20 to 16, while the comparison group had worsened in risk from 18 to 21. A significant effect of the covariate age, $F(1, 88) = 12.14$, $p = .001$, was found and the raw data indicated that older participants' risk scores were reducing marginally more than younger participants. This should be viewed with some caution, however, as the older participants generally had higher LSI-R scores at the start of intervention and therefore had more room to improve. This was to be expected given the well-known findings from cohort studies across the globe (see, for instance, Farrington, Jolliffe, Loeber, & Homish, 2011) indicating that male offending behaviour increases during the teenage years, peaking at around 17. A favourable effect size was noted ($\eta^2 = .19$), which compares well with previous studies on the impact of interventions on adult recidivism. Effect sizes for the majority of these have ranged from 0.12 to 0.13 (e.g., Gottschalk, Davidson, Gensheimer, & Mayer, 1987; Whitehead & Lab, 1989). In fact, Lipsey's (1995) meta-analysis of 443 studies of juvenile offenders yielded treatment effects of between 0.05 and 0.08. Stronger treatment effects have been seen in adult offenders (McGuire, 2002) but these still range from 0.10 to 0.29.

Table 1. Comparison of LSI-R Subscores at Times 1 and 2 for the Intervention Group Paired Samples *t* Test

	Average scores			
	Time 1	Time 2	*t*	Significance (one-tailed; Bonferroni correction, $p = .004$)
Constructive recreation	1.7	1.0	6.2	<.001
Attitudes/engagement with education	2.5	1.6	6.5	<.001
Criminal attitudes/orientation	2.1	1.2	4.9	<.001
Family problems	2.1	1.7	3.4	.001
Emotional problems	1.0	0.7	3.7	.001
Financial problems	0.6	0.4	3.2	.002
Companions	2.5	2.4	1.0	.333, *ns*
Accommodation	1.3	1.1	1.0	.311, *ns*
Drugs/alcohol	1.8	1.6	1.4	.162, *ns*
Criminal history	2.8	2.9	−1.9	.06[a]

Note: LSI-R = Level of Service Inventory–Revised.
[a]Significant worsening in this static factor but is actually a reflection of increased disclosure of previous offences over time.

Further analysis of the subcomponents of the LSI-R indicated six areas where statistically significant improvements were taking place, as Table 1 indicates.

Significant reductions in risk were seen in key subcomponents associated with leisure and recreational activities, attitudes and engagement with education, and anticriminal attitudes and orientation. These were followed by significant improvements in family functioning, financial problems, and emotional problems. It is also notable that this is occurring at a much earlier age than is typical for most offenders.

Areas where less change appeared to be occurring in the current evaluation of PYOP's participants were in relation to antisocial peers and drugs/alcohol problems. The findings of the ongoing evaluations have been useful to practitioners in refining the input to individual participants.

A univariate ANOVA conducted on participants with PYOP for a year or more (Group 1a, $n = 46$) demonstrated a further statistically significant drop in risk scores at Time 3 (1 year), $F(2, 45) = 32.1$, $p < .001$. Pairwise comparisons showed that the significant change had occurred between Times 1 and 2 (after 6 months), and Times 1 and 3 (after 1 year). Although the drop in risk scores between Times 2 and 3 did not reach statistical significance, Figure 2 indicates that improvements made by Time 2 were sustained and built on over the following 6 months. Moreover, more than half (54%) of this group were aged 13 or below (as opposed to 43% in Group 1 above). Two further subcomponents of the LSI-R reached statistical significance in this analysis: a decrease in accommodation problems (highlighting the broad reach of PYOP to the criminogenic needs of the whole family) and very importantly, lower association with anticriminal friends and associates.

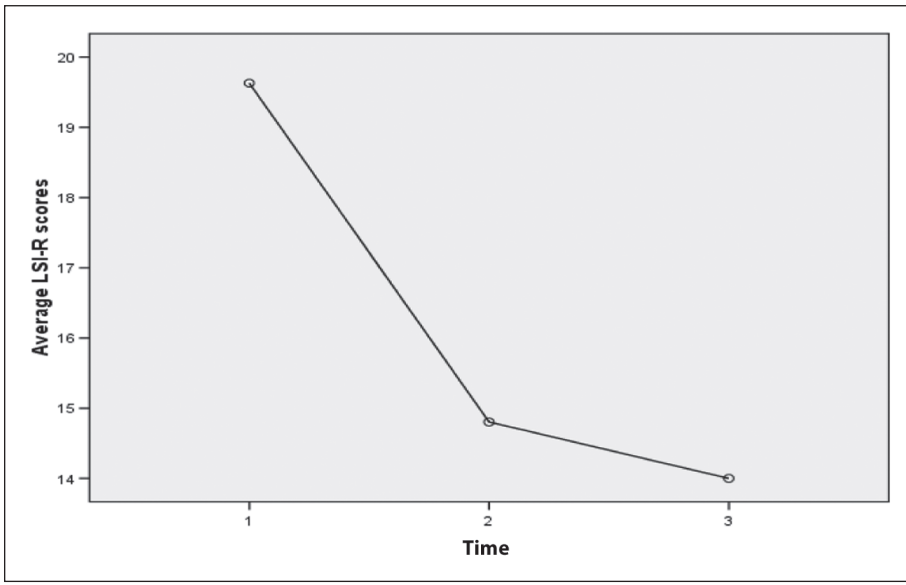

Figure 2. LSI-R improvements for Group 1a (n = 46) at Times 1 to 3 (after 1 year at PYOP)
Note: LSI-R = Level of Service Inventory–Revised; PYOP = Preventing Youth Offending Program.

Group 1b constituted 24 participants who had been with PYOP for 2 years or more and included some very challenging cases. Only 4 were aged above 12. It is to the credit of the program that they can engage such young participants for the long term. We know from studies of attrition that it is the youngest participants who are most likely to drop out (McMurran & McCulloch, 2007). Analysis of this group indicated that PYOP had managed to maintain a statistically significant impact over this time, $F(4, 23) = 5.3$, $p = .001$, but pairwise comparisons showed no statistical significance between Times 1 and 5 (after 2 years of intervention) as Figure 3 shows.

This pattern can be explained by the fact that the long-termers are the neediest and most complex group and are highly likely to show variations over such a sustained length of time. However, it is notable that these more ambiguous findings are the result of only 4 out of the 24 participants whose risk increased over time, with 2 in custody shortly after the 2-year milestone. Improved subcomponents on the LSI-R for the group overall were identical to Group 1a, though interestingly "anticriminal attitudes" achieved only marginal significance ($t = 2.04$, $p = .55$). It is important to remain mindful that with this particularly testing and vulnerable group, the overall impact of PYOP remained positive.

On a methodological note, it could be argued that longer term participants in Groups 1a and 1b may have been retained because they were doing well or because they were

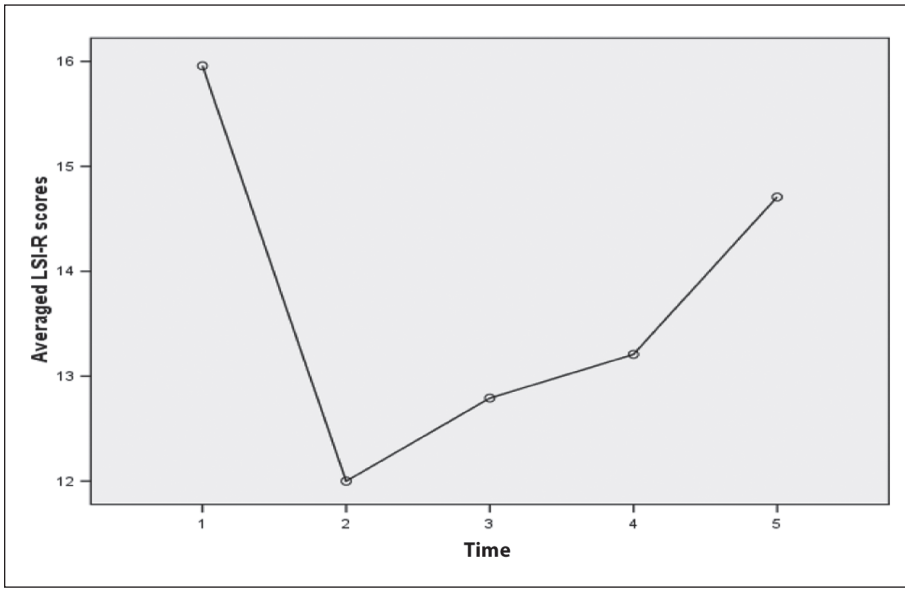

Figure 3. LSI-R improvements for Group 1b (n = 24) at Times 1 to 5 (after 2 years at PYOP)
Note: LSI-R = Level of Service Inventory–Revised; PYOP = Preventing Youth Offending Program.

easier to engage. However, there was no indication of this in either their case notes or their LSI-R scores. For instance, for those who had only one review in Group 1, their mean LSI-R score at Time 1 was 20, compared with 23 for those in Group 1a and 16 in Group 1b, $F(2, 66) = 3.97$ $p = .024$. Although the LSI-R mean scores go up and down, they do not suggest that the majority of the longer term clients had fewer needs. Furthermore, evidence suggests that the participants who were continuing with the intervention were not doing so purely because they were those who had not dropped out. We coded LSI-R case notes for the 1st year of intervention. For those who exited the program at some time between 6 and 12 months (n = 21), notes indicated that only 3 dropped out through disengagement with the project. The remainder left because the coordinators (and participants) agreed that they had come to the end of their program, and intervention was no longer needed—in other words a successful termination. Figure 3 also indicates that Group 1b became more challenging as intervention progressed. They were not, therefore, the soft option regarding treatment.

Police Charges Analysis (YOIS). Police charges data for the 6 months before PYOP intervention, and at 6 monthly intervals subsequently, was collected as an additional, proxy measure for reoffending for the 55 members of Group 1 above the age of criminal responsibility in the United Kingdom (10 years) and all 24 members of the comparison group. As noted above, these data are less reliable than the LSI-R risk scores. For

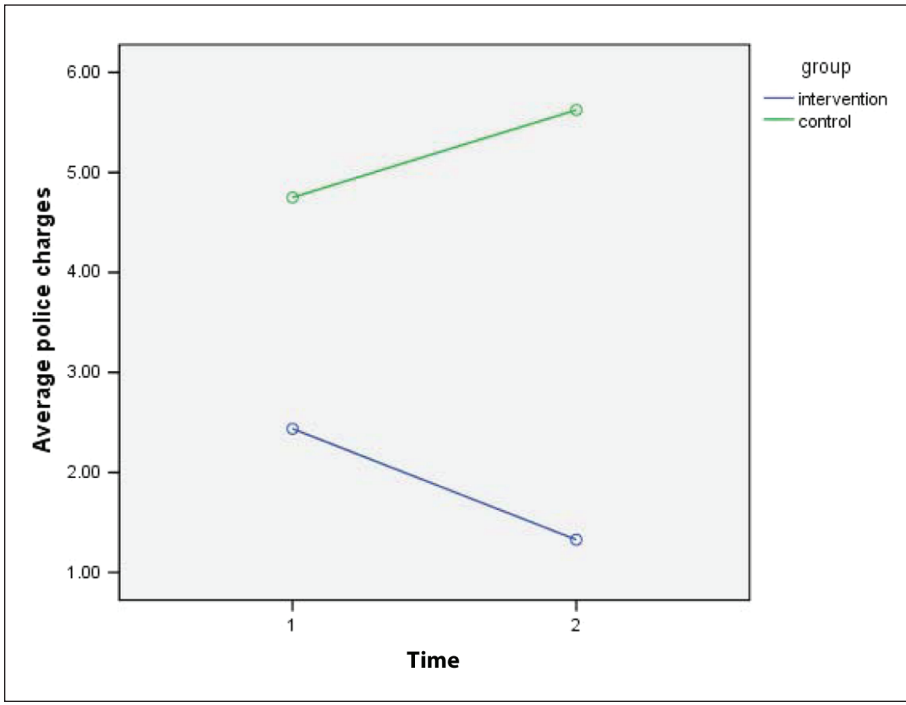

Figure 4. Average number of police charges for 6 months before PYOP intervention (Time 1) and first 6 months during intervention (Time 2) for intervention group and comparison group
Note: PYOP = Preventing Youth Offending Program.

instance, a number of members of both the intervention group and comparison group had no charges in the 6 months before PYOP (even though they were known to be offending and disclosed offences in the first LSI-R assessment), but several in the 6 months previous to that. These data nevertheless furnish us with a reasonable secondary indication of background offending behaviour. A repeated measures ANOVA indicated a significant time by group interaction, $F(1, 77) = 4.68$, $p = .03$, $\eta^2 = .57$, indicating a significant difference between the groups over time. Post hoc comparisons, however, indicated a less clear picture than in the risk analysis above. Although the two groups were not significantly different regarding the charges they received in the 6 months before PYOP intervention, $t(1,79) = -1.9$, $p = .06$, the test did approach significance with average numbers of charges lower for the intervention group ($M = 2.47$, $SD = 3.15$) than the comparison group ($M = 4.33$, $SD = 5.57$). By Time 2, however, the comparison group had worsened significantly, whereas the intervention group had improved, as Figure 4 indicates. During the first 6 months of intervention, the average number of charges against the PYOP group had dropped to 1.32 ($SD = 1.99$), whereas it was 5.7 ($SD = 7.5$) for the comparison group, $t(1,77) = -2.77$, $p < .0001$ with Bonferroni correction.

Notwithstanding the less reliable nature of the data, the overall picture supports the LSI-R risk findings described above.

The 37 participants aged above 10 from Group 1a (1+ years of intervention) showed a clear and statistically significant drop in police charges from Time 1 to Time 3, $F(1, 37) = 4.80$, $p = .03$, from 1.7 on average to 0.8, and Group 1b ($n = 17$) reduced their mean number of charges from 1.1 in the 6 months before PYOP to 0.5 after 2 years of intervention, suggesting a sustained improvement with this complex group.

Discussion

Results suggest that PYOP intervention had a favourable impact on reducing the risk of reoffending as measured by the LSI-R in this very young group of offenders, with very favourable effect sizes. This is good news given the average LSI-R score at Time 1 for Group 1 was 20—This is the average score for adult male offenders in England and Wales (Raynor, 2007). The clearest impact was seen in the first 6 months of intervention, after which the reduction in risk was sustained and further improved on. This is supported by a notable drop in police charges against the intervention group during treatment in contrast to an increase for the comparison group.

Results were also encouraging regarding the subcomponents of change. Areas of significant change for Group 1 as measured by the LSI-R included leisure and recreational activities (such as participation in a pro-social organised activity and good use of free time), attitudes and engagement with education (reflecting engagement with schoolwork, peers, and authority figures while in education), anticriminal attitudes and orientation, and emotional problems. This suggests improvement in three of the "big four" elements of criminality according to Andrews and Bonta's (1995, 2010) theory of criminal conduct: pro-criminal attitudes, associates, antisocial personality, and criminal history (the fourth element being a historical variable and therefore not open to change). Improvements in family functioning and financial problems highlight the project's holistic, multimodal approach to offenders' and their families' needs. It is also impressive that this intervention is addressing these problems at a much earlier age than is typical for most offenders. This approach could mean reduced costs to the criminal justice system and improved quality of life for these children and their families. Analysis of the 46 young people who had received 1 year or more of intervention (Group 1a) indicated further improvement at Time 3 in both LSI-R scores and offending behaviour, indicating the importance of long-term intervention and responding to changing needs.

Analysis of the long-term intervention group (Group 1b) highlighted the extra challenges associated with this considerably needy and especially young population (modal/median age = 11). It revealed an interesting finding that, although statistically significant change was attained overall, only marginally significant improvements were achieved in relation to a key subcomponent "criminal attitudes and orientation." This surprising finding may signify how early problematic attitudes and orientation start to form in the core of offenders who, if not rehabilitated, may present the biggest challenge to the criminal justice system in the long term. It underlines the importance

of addressing criminal attitudes from a very early age. It may also indicate that long term, these attitudes are the hardest nut to crack. However, to our knowledge, no other evaluations of the subcomponents of criminal conduct exist for this age group and clearly more are needed. The long-term nature of the intervention and its ability to engage such young participants is to be commended and may be a result of the specific responsivity principle that seems inherent in the program (see below). The single other program for children of this age group known to be offending that has published positive evidence not only offers 12 weeks of intervention to each child (Koegl et al., 2008) as a core program but also offers a variety of add-on interventions depending on the needs of the child and his or her family. It is of note too, that in the only published program evaluation known to have success with psychopathic teenage boys, improvements in terms of reduced violent reoffending and improved compliance were associated with length of time in treatment (Caldwell, McCormick, Umstead, & Van Rybroek, 2007). Participants were engaged in highly structured residential treatment for an average of 45 weeks. These findings, coupled with ours above, underline the importance of a flexible approach to dosage, and maintaining intervention for as long as needs exist, especially with this vulnerable age group. This flexibility is virtually unheard of with regard to adult interventions.

The favourable results from the ongoing summative evaluation of PYOP provided a rationale to look at in-program features, such as treatment types and intensity, in other words, the how and the why of the program's impact (Wormith et al., 2007) in line with the overall aims of the article.

Study 2—Analysis of Treatment Type and Intensity

Aims

In this study, we aimed to understand the nature of the program more clearly, that is, what types of intervention were involved, whether any types of intervention were more common or more effective than others, and to see whether treatment dosage had an effect on outcome. It was in this analysis that issues associated with specific responsivity in relation to the tailoring of treatment to individual needs emerged more clearly.

Method

Design. This study involved a correlational design to see whether there was an association between particular treatments and reduction in risk and/or level of dosage and reduction in risk.

Participants. At the time of data collection for this study, there were 39 participants in the intervention group[3] (Group 1) and 21 participants in the comparison group. Data were coded for all 60 participants. All but 1 of Group 1 was male as were all the

comparison group members. Ages ranged in Group 1 from 7 to 16 ($M = 13$), and 12 to 16 in the comparison group ($M = 14$).

Procedure. The personal files held by PYOP for each member of the intervention group and comparison group were examined for intervention type and dosage, which was coded for analysis. These were recorded for the 1st year of intervention, but for the purposes of this article, we focus on the first 6 months as the raw data were more complete and the greatest impact of the intervention was seen during this period. Four broad types of intervention became apparent as data were coded from the files: one-to-one support, educational support, group work, and outdoor activities, and the elements of these are described below.

One-to-one support. This included one-to-one constructive activities between the assigned project worker designed to engage the young person in prosocial activity, for example, sports, walking, playing pool, lunch, or shopping for essential items. Mentors were trained to encourage participants to discuss any problems or issues arising, to encourage participants to generate pro-social solutions, and to challenge inappropriate or unacceptable attitudes or behaviours (such as not attending school or fighting). This also included home visits when it was not logistically possible to organise a meeting elsewhere or when the young person needed extra support. Project workers would also feed back to the coordinator if they became aware that particular needs had arisen in the family in general and this would then result in a family visit by the coordinator to facilitate support. Unfortunately, the recording of these visits was very unreliable so only visits with the young people have been coded.

Educational support. For the participants most at risk of school exclusion and those who were in the process of being reintegrated to mainstream education after exclusion, educational support was given. This usually involved the project worker picking up the child from home, ensuring they had breakfast, and accompanying them to school (or pupil referral units for those excluded) to offer support in lessons, for a number of mornings a week. Visits to museums and libraries were also built in to enrich the child's experience if not in full-time mainstream school.

Group work. PYOP offered 12-week challenging offending behaviour and challenging antisocial behaviour groups. These were based on cognitive-behavioural principles and involved problem solving and social perspective taking, challenging attitudes to offending and antisocial behaviour, criminogenic peers, constructive use of time, and motivation. Other work in groups involved art, music, and drama workshops either bought in by PYOP or offered by the project workers.

Outdoor activities. Twelve-week outdoor activity programs were offered to some participants. Each week the participants would take part in a different activity. This was designed to build teamwork, confidence, trust, and life skills. In addition, other 1-day outdoor activity sessions were delivered.

All of the above activities were planned at high-risk times to divert the young person from possible offending or mixing with criminogenic peers. All dosage was recorded in hours.

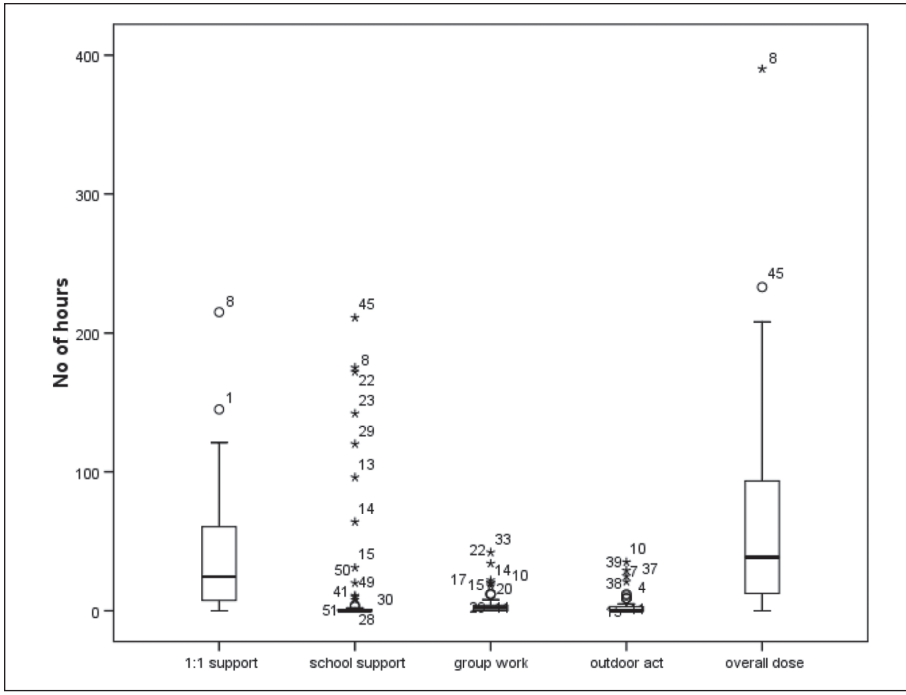

Figure 5. Box plot indicating range of hours that individual participants received by intervention type
Note: Numbers on the chart indicate cases.

To assess the relative contribution of the various intervention types, we considered conducting a multiple regression using the number of hours of the four intervention types as predictor variables and amount of risk reduction (LSI-R score) as the dependent variable. However, this was not possible as all the predictor variables were markedly positively skewed and the absolute number of hours was too small for two of the variables. We could not rectify the problem of skewed distributions by using a transformation as there was a relatively large number of zero values in the predictor variables. Therefore, to explore the impact of the program, we present simple bivariate nonparametric correlations (Spearman's rho) for the intervention group between the treatment hours of the different kinds of treatment and the amount of overall improvement (reduction) in risk score.

Results

The most striking feature of the treatment data is the enormous range of input that different cases received, as Figure 5 indicates. Participants received between 5 and 390 hr of intervention during their first 6 months at PYOP, depending on what was

Table 2. Relative Contribution of Different Types of Treatment

Treatment	Total number of hours	Percentage total input
One to one	1,955	61
Educational support	817	26
Group work	259	8
Outdoor activity	179	5

Note: Unequal distribution of hours across different types of activity, $\chi^2(3) = 50.12$, $p < .0001$.

deemed appropriate by project managers. Interestingly, there was no correlation between treatment intensity and level of risk at first assessment suggesting that input was being matched to the level of need using other indices (see "Discussion" section of Study 2 for examples). Figure 5 highlights the extremely high level of specific responsivity in the program, which is a key element of PYOP.

A negative correlation was found between age and treatment intensity, $\rho = -.405$ (39), $p = .01$, indicating younger participants were receiving greater input.

Table 2 indicates the relative contribution of different types of treatment. It is clear that the allocation of treatment hours across individual clients is very positively skewed with a number of extreme outliers. Client 8, for instance, was allocated 20% of the total therapy hours.

No association emerged between any one type of treatment and overall improvement in risk score but, as expected, a significant and relatively strong negative correlation could be seen between reduction in risk and total dosage, $\rho = -.42$ (39), $p = 004$, suggesting that no one intervention type was more effective than another and that the package as a whole was working. It also suggests that staff judgement regarding relative input of different types of intervention was effective and this needs further investigation.

Discussion

We believe that the ability of the program to engage such young and clearly challenging participants for the long term is evidence of the appropriateness of the interventions and the quality of the therapeutic alliance between clients and staff. It could also be seen as a rare example of the general responsivity principle with this age group. Although we did not take direct measures of quality, we know from interview and observation that the three key components from Lowencamp, Latessa, and Smith's (2006) review of quality in adult corrections—program implementation, offender assessment, and program evaluation—were all strongly upheld. Indeed, Lowencamp et al.'s emphasis on the qualifications and experience of program leaders and their overall direct involvement in the program, including recruitment of high-quality staff, was strongly reflected in PYOP.

Furthermore, the notable variation of input to different child participants suggests an underlying commitment to specific responsivity in the program and may well be one of the key elements to the apparent success of the program. Constraints on the research remit did not allow us to examine the exact decision-making process regarding how individual clients were allocated to specific interventions but nevertheless there is clear evidence that each child received a very individualised program based on need. Treatment intensity was not associated with risk score, but was associated with age (younger participants receiving more input). PYOP staff indicated that for all participants, and especially the very young ones, their treatment decisions were driven by need (including noncriminogenic need) rather than risk. Need was assessed not only by the LSI-R but also by using a variety of other social-work-based assessments (Department of Health, 1988; Onset, 2006); the experience, discretion, and knowledge of the project coordinator[4]; and the specific strengths and interests identified in the participant. At the start of intervention, treatment decisions were greatly influenced by what the participant themselves found acceptable (which sits most closely with the approach of the GLM; Ward et al., 2012). Staff felt this was important in empowering the young people to make decisions about their own lives as most felt powerless at the point of entry. As trust was built up, participants were more accepting of further intervention negotiated by project workers.

Ostensibly, this approach to intervention may appear to contravene the risk principle in adult offenders (i.e., of matching risk and need with level of service), but this highlights the importance of conducting research with this qualitatively different age group. An anomaly exists when assessing very young offenders, which complicates the scene. Risk is obviously heavily associated with the passing of time and life experience, and therefore the very young participants will have lower risk scores by default. In this case, we should be a little more cautious about interpreting LSI-R scores for very young participants until we have clearer knowledge about risk in this group. This aside, the findings regarding risk and need above may simply highlight a difference in the balance and nature of intervention required for this age group.

In terms of the nature of intervention, one-to-one support and educational support were the most frequently used, but, as noted, no single treatment was strongly associated with reduction in LSI-R scores on its own. Instead overall dosage, whatever mix of intervention that consisted of, was significantly associated with improvement (in line with Lipsey's meta-analysis on young offenders in 1995). This may indicate that the integrated intervention package was working as a coherent whole and was involving the right mix of interventions for individual participants, and decisions about level and type of treatment seemed to be effective.

Limitations

It is important to note that, despite the positive findings, there were several limitations to the present studies that make the results suggestive rather than indicative, and further

research is clearly needed. The sample sizes of both studies were very small at $n = 67$ in Study 1 and $n = 39$ in Study 2. As noted above, offenders of this age are rare and because of the small, intensive nature of the program, it has taken some time for treated participants to accrue. Similarly, the comparison group was very small ($n = 24$) and ad hoc in nature, albeit meeting the standards of the "incidental comparison group" (Marshall & McGuire, 2003). Despite their rarity, it is very important that we continue to research this vulnerable group, and we hope that over time, more evaluations of interventions with this age group will emerge, so that we can evaluate our findings in relation to other interventions and contribute further to our understanding of childhood criminality.

To keep the costs of the evaluation down for PYOP, it was agreed that LSI-R assessments for both the treatment group and comparison group were undertaken by the project coordinator rather than by an independent researcher. Assessments were 6 months apart and there was evidence of increasing and decreasing risks in both the treated group and comparison group at Time 2, so we were reasonably confident that assessments were sufficiently free from bias. Future research should ideally aim to employ independent researchers for data collection purposes, however.

The risk assessment tool used to measure changes in risk in Study 1 was developed for adult male offenders rather than the children and adolescents in our study. At the time of embarking on the evaluation, no risk assessment tools existed for children and the decision to use one of the most robust risk measures for adults was taken. The LSI-R is also based on components empirically linked with the onset of criminality, many of which have a strong impact during childhood and adolescence. As noted above, youth versions of the LSI-R have been published in recent years, but these are more therapeutic and case management oriented than risk oriented and it was also important methodologically that we continued with the same measure we had begun the evaluation with, for comparison reasons. Until the Youth Level of Service Inventory (YLSI), for example, is tested for its risk prediction properties, research in this field may have to rely on adult-based tools, though future research could ideally incorporate a broader range of assessment tools.

It was beyond the scope of this research to describe in detail the approach to case formulation used by staff at PYOP, as this was not part of the original research remit. However, now that we have data suggesting that what staff are doing seems effective, our next step is to examine the decision-making process regarding how staff allocate types and levels of input for participants and produce illustrative case studies that would aim to inform practice more directly. We are confident that responsive decisions are being made. Now, we need to observe and understand the nature of the process.

General Discussion

Our analysis of PYOP has produced some interesting findings regarding this age group of offenders: a population that we know so little about in terms of effective intervention. We have shown that, in terms of general principles, the RNR approach to intervention sits quite well with this unusually young population of offenders. By

analysing the nature of PYOP, we have uncovered two unusual practitioner-led features that are worthy of further consideration with this age group, and for which more research is clearly needed. First, an extremely high level of responsivity in the program was apparent, both general (in terms of the multimodal and cognitive-behavioural approach) and specific (in terms of addressing each individual's need profile as closely as possible and for as long as the child needs it). Second, and relatedly, treatment intensity was matched not only to risk but also to both criminogenic and noncriminogenic needs. With such a vulnerable age group, it is difficult to imagine not having to address need in a more holistic way, and further RNR-style evaluations of programs for this age group are needed to confirm whether this slightly adjusted configuration of the RNR model holds for very young offenders.

PYOP's approach also aligns well with findings from Lipsey's (2009) recent meta-analysis of effective interventions for juvenile offenders (aged 12-21) in which strongly therapeutic approaches, aimed at high-risk juveniles, were the most successful. We encourage more evaluations such as ours, as we feel that our evidence begs for further investigation if we are to begin to understand what is effective with this age group. Indeed, as Wormith et al. (2007) recommended, we need to look ever more closely at the process and nuances of change as well as reduction in risk. In terms of Wormith et al.'s review of the current and future landscape, PYOP fits best within the newer generation of more client responsive, strengths-oriented approaches to intervention such as the GLM with its explicit aim to address each unique profile of offender's needs to increase chances of engagement. Although we are in the early days of evaluation, evidence is beginning to accrue with adults (e.g., Lindsay, Ward, Morgan, & Wilson, 2007; Mann, Webster, Schofield, & Marshall, 2004) that this approach enhances the outcomes of intervention in comparison with the more traditional risk-oriented models. With younger offenders also, there are some studies that support this view. Veira et al. (2009), in a study of the treatment of 122 young offenders aged 12 to 18, found that the greater the number of criminogenic needs met and individual responsivity issues taken into account, the lower the frequency of reoffending and the longer the length of time an offender took between treatment and offending. Similarly, Koegl et al. (2008) found a positive correlation with the number of treatment components administered and the level of improvement in Child Behaviour Checklist (Achenbach, 1991) that measures delinquency as one of its components in a group of 66 offending children aged 6 to 11 years. Our evaluation here will add to the growing evidence for a strength-based approach, such as the GLM (Ward et al., 2012).

In adult correctional policy, assessment of risk has traditionally been the first priority and there is evidence, as we noted in the "Introduction," that responsivity is often a last priority. However, Andrews and Bonta (2010) have clearly shown that in the most successful programs for adults, all three RNR principles are given equal importance, with an incremental drop in recidivism the closer the principles are adhered to. Other examples of an increased focus on responsivity can be seen in a number of relatively new treatments for adult offenders. As well as those mentioned in the "Introduction," the incorporation of pretreatment motivational interviewing (Miller &

Rollnick, 2002) and one-to-one orientation sessions, plus increased use of one-to-one sessions during treatment (see Nee & Farman, 2008, for work on dialectical behaviour therapy for women prisoners) pay heed to the need to underpin group work with individually tailored support for clients' needs. These features show great promise in enhancing the efficacy of offender treatment (McMurran, 2009; Nee & Farman, 2008), and in the present study, it was the use of one-to-one support with everyday, pro-social activities that was used most frequently and with the widest range of participants.

There is an obvious and sound need to monitor program integrity in any intervention and an ethical responsibility on the part of policy makers to ensure offenders get a reasonably standardised, evidence-based intervention. But in the recent past, this consideration has overridden the need for flexibility for a program to be truly responsive to the needs of the individual (and consequently have greater chances of facilitating change; Bonta, 1995; Ellis & Winstone, 2002). It is encouraging that the flexibility message now seems to be getting through to policy makers, and the newer generation of interventions noted by Wormith et al. (2007) are underpinned by the risk principle but also acknowledge and accommodate the need principle and responsivity principle more fully. Indeed, both Blanchette (2002) and Ward and Stewart (2003) noted the importance of incorporating noncriminogenic needs in the treatment of female offenders and sex offenders, respectively. The same seems to be valid for the very youngest of offenders.

The evidence above suggests that a PYOP-type approach to intervention may be effective in rehabilitating and arresting the development of criminality in particularly challenging preteenage offenders. The long-term economic payoff alone in terms of crime prevented (Welsh & Farrington, 2011) and the ramifications this has for the criminal justice system should be enough to convince policy makers to invest in further research in the first instance. It is important to consider this too within Andrews and Bonta's (2010) recent and gloomy prediction of increased incarceration rates and reductions in rehabilitation in the U.S. criminal justice system. In short, for ethical as well as economic reasons, it is time to do a better job of including and engaging with the youngest, the poorest, the most marginalised, the highest risk, and, in other words, the neediest offender populations.

Acknowledgment

The authors would like to thank Professor Aldert Vrij for his support throughout the publication process, Professor James McGuire for his comments on an early draft, and Rhona Lucas, Bruce Marr, and the participants of Preventing Youth Offending Program (PYOP).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by the Portsmouth City Council to undertake an independent evaluation of Preventing Youth Offending Program (PYOP).

Notes

1. Though it should be noted that a preexisting and widely used program (Think First accredited in the United Kingdom since 1996) has 10 individual sessions and uses a goal-orientated approach.
2. Such as child skills training (Losel & Beelman, 2003).
3. These data were collected 18 months prior to Study 1.
4. Dowden and Andrews (2004) have also noted the importance of the skills and experience of those delivering treatment in the successful rehabilitation of the offender.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington: Department of Psychiatry, University of Vermont.

Andrews, D. A., & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto, Ontario, Canada: Multi-Health Systems.

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating criminal justice policy and practice. *Criminal Justice and Behavior*, *16*, 39-55.

Andrews, D. A., Bonta, J., & Wormith, S. (2006). The recent past and near future of risk/need assessment. *Crime & Delinquency*, *52*, 7-27.

Blanchette, K. (2002). Classifying female offenders for effective intervention: Application of the case-based principles of risk and need. *Forum on Corrections Research*, *14*, 31-35.

Bonta, J. (1995). The responsivity principle and offender rehabilitation. *Forum of Corrections Research*, *7*, 34-37.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior*, *29*, 355-379.

Bonta, J., & Andrews, D. A. (2007). *Risk-need-responsivity model for offender assessment and rehabilitation* (User Report 2006-2007). Ottawa, Ontario: Public Safety Canada.

Burrowes, N., & Needs, A. (2009). Time to contemplate change? A framework for assessing readiness to change in offenders. *Aggression and Violent Behavior*, *14*, 39-49.

Butler, S., Fearon, P., Atkinson, L., & Parker, K. (2007). Testing an interactive model of symptom severity in conduct disordered youth: Family relationships, antisocial cognitions, and social-contextual risk. *Criminal Justice and Behavior*, *34*, 721-738.

Caldwell, M. F., McCormick, D. J., Umstead, D., & Van Rybroek, G. J. (2007). Evidence of treatment progress and therapeutic outcomes among adolescents with psychopathic features. *Criminal Justice and Behavior*, *34*, 573-589.

Cann, J. (2006). *Cognitive skills programmes: impact on reducing reconviction among a sample of female prisoners* (Home Office Research Findings No. 276). London, England: Home Office.

Department of Health. (1988). *Protecting children: a guide for social workers undertaking a comprehensive assessment*. London, England: HMSO.

Dowden, C., & Andrews, D. A. (2004). The importance of staff practices in delivering effective correctional treatment: A meta-analysis of core correctional practices. *International Journal of Offender Therapy and Comparative Criminology*, *48*, 208-214.

Ellis, T., & Winstone, J. (2002). The policy impact of a survey of programme evaluations in England and Wales: Towards a new corrections-industrial complex? In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies*. Chichester, UK: Wiley.

Falshaw, L., Friendship, C., Travers, R., & Nugent, F. (2003). *Searching for "What Works": An evaluation of cognitive skills programmes* (Home Office Research Findings 206). London, England: Home Office.

Farrington, D.P. (2002). Developmental criminology and risk-focused prevention. In M. Maguire, R. Morgan, & R. Reiner (Eds.), *The Oxford handbook of criminology* (3rd ed.). Oxford, UK: Oxford University Press.

Farrington, D. P., & Jolliffe, D. (2002). *A feasibility study into using a randomised controlled trial to evaluate treatment pilots at HMP Whitemoor* (Home Office Online Report 14/02). London, England: Home Office.

Farrington, D. P., Jolliffe, D., Loeber, R., & Homish, D. L. (2011). Scaling up from juvenile court records to offenses committed. In A. R. Roberts & D. W. Springer (Eds.), *Juvenile delinquency and juvenile justice*. Sudbury, MA: Jones and Bartlett.

Farrington, D. P., & Welsh, B. (2007). *Saving children from a life of crime*. New York, NY: Oxford University Press.

Garrido, V., & Morales, L. (2007). *Serious (violent and chronic) juvenile offenders: A systematic review of treatment effectiveness in secure corrections*. Oslo, Norway: Campbell Collaboration Crime and Justice Group.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of adult offender recidivism: What works! *Criminology*, *34*, 575-607.

Gottschalk, R., Davidson, W. S., II, Gensheimer, L. K., & Mayer, J. P. (1987). Community based interventions. In H. C. Quay (Ed.), *Handbook of juvenile delinquency* (pp. 266-289). New York, NY: Wiley.

Hollin, C. R. (2002). Risk-needs assessment and allocation to offender programmes. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies to reduce re-offending*. Chichester, UK: Wiley.

Hollin, C. R. (2006). Offending behaviour programmes and contention: Evidence-based practice, manuals, and programme evaluation. In C. R. Hollin & E. J. Palmer (Eds.), *Offending behaviour programmes: Development, application, and controversies*. Chichester, UK: Wiley.

Hollin, C. R. (2008). Evaluating offending behaviour programmes: Does only randomization glister? *Criminology and Criminal Justice*, *8*, 89-108.

Hollin, C. R., & Palmer, E. J. (2006). Criminogenic need and women offenders: A critique of the literature. *Legal and Criminological Psychology*, *11*, 179-195.

Hollin, C. R., Palmer, E. J., & Clark, D. (2003). The Level of Service Inventory–Revised profile of English prisoners: A needs analysis. *Criminal Justice and Behavior*, *30*, 422-442.

Home Office. (1996). *Guidance for the probation service on the Offender Group Reconviction Scale* (Probation Circular 63/1996). London, England: Author.

Hsu, C., Caputi, P., & Byrne, M. (2009). The Level of Service Inventory–Revised (LSI-R): A useful risk assessment measure for Australian Offenders. *Criminal Justice and Behavior*, *36*, 728-740.

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base and improve patient care. *American Psychologist*, *63*, 146-159.

Koegl, C., Farrington, D., Augimeri, L., & Day, D. (2008). Evaluation of a targeted cognitive-behavioural programme for children with conduct problems—The SNAP® Under 12 Outreach Project: Service intensity, age and gender effects on short- and long-term outcomes. *Clinical Child Psychology and Psychiatry*, *13*, 419-434.

Langlands, R, Ward, T., & Gilchrist, L. (2009). Applying the Good Lives Model to male perpetrators of domestic violence. *Behaviour Change*, *26*, 113-129.

Lindsay, W. R., Ward, T., Morgan, T., & Wilson, I. (2007). Self-regulation of sex offending, future pathways and the Good Lives Model: Applications and problems. *Journal of Sexual Aggression*, *13*, 37-50.

Lipsey, M. W. (1995). Can rehabilitative programs reduce the recidivism of juvenile delinquents? In J. McGuire (Ed.), *What works: Reducing reoffending. Guidelines from research and practice*. Chichester, UK: Wiley.

Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A metaanalytic overview. *Victims & Offenders*, *4*, 124-147.

Lipsey, M. W, Landenberger, N., & Wilson, S. (2007). *Effects of cognitive-behavioral programs for criminal offenders*. Oslo, Norway: Campbell Collaboration Crime and Justice Group.

Lipsey, M. W., & Wilson, D. B. (1998). Effective intervention for serious juvenile offenders: A synthesis of research. In R. Loeber & D. P. Farrington (Eds.), *Serious & violent juvenile offenders: Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant disorder and conduct disorder: The last ten years, part 1. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 1468-1482.

Loeber, R., & Farrington, D. P. (Eds.). (2001). *Child delinquents: Development, interventions and service needs*. Thousand Oaks, CA: Sage.

Losel, F., & Beelman, A. (2003). Effects of child skills training in preventing antisocial behaviour: A systematic review of randomized evaluations. *Annals of the American Academy of Political and Social Science*, *84*, 1-27.

Lowencamp, C. T., Latessa, E., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy*, *5*, 201-220.

Lynam, D. R., Charnigo, R., Moffitt, T. E., Raine, A., Loeber, R., & Stouthamer-Loeber, M. (2009).The stability of psychopathy across adolescence. *Development and Psychopathology*, *21*, 1133-1153.

Mann, R. E., Webster, S. D., Schofield, C., & Marshall, W. L. (2004). Approach versus avoidance goals in relapse prevention with sexual offenders. *Sex Abuse*, *16*(1), 65-75.

Marshall, W. L., & McGuire, J. (2003). Effect sizes in the treatment of sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, *47*, 653-663.

McGuire, J. (2002). Criminal sanctions versus psychologically-based interventions with offenders: A comparative empirical analysis. *Psychology, Crime & Law*, *8*, 183-208.

McGuire, J. (2006). General offending behaviour programmes: Concept, theory, and practice. In C. R. Hollin & E. J. Palmer (Eds.), *Offending behaviour programmes: Development, application, and controversies*. Chichester, UK: Wiley.

McGuire, M. (2008). A review of effective interventions for aggression and violence. *Philosophical Transactions of the Royal Society*, *363*, 2577-2597.

McMurran, M. (2009). Motivational interviewing with offenders: A systematic review. *Legal and Criminological Psychology*, *14*, 83-100.

McMurran, M., & McCulloch, A. (2007). Why don't offenders complete treatment? Prisoners' reasons for non-completion of a cognitive skills programme. *Psychology, Crime & Law*, *13*, 333-343.

Miller, W. R., & Rollnick, S. (2002). *Motivational interviewing: Preparing people for change*. New York, NY: Guilford.

Moffitt, T. E. (1993). Life-course-persistent and adolescence-limited antisocial behaviour: A developmental taxonomy. *Psychological Review*, *100*, 674-701.

Nee, C., & Ellis, T. (2005). Treating offending children: What works? *Legal and Criminological Psychology*, *10*, 133-148.

Nee, C., & Farman, S. (2008). Treatment of borderline personality disorder in prisons: Findings from the two dialectical behaviour therapy pilots in the UK. In J. C. Hagen & E. I. Jensen (Eds.), *Personality disorders: New research*. New York, NY: Nova Science Publishers.

Odgers, C., Caspi, A., Broadbent, J., Dickson, N., Hancox, R., Harrington, H., & Moffitt, T. (2007). Prediction of differential adult health burden by conduct problem subtypes in males. *Archives of General Psychiatry*, *64*, 476-484.

Onset. (2006). *Centre for criminology*. Oxford, UK: Oxford University Press.

Palmer, E. J., McGuire, J., Hatcher, R. M., Hounsome, J. C., Bilby, C. A. L., & Hollin, C. R. (2008). The importance of appropriate allocation to offending behaviour programs. *International Journal of Offender Therapy and Comparative Criminology*, *52*, 206-221.

Raynor, P. (2007). Risk and need assessment in British probation: The contribution of LSI-R. *Psychology Crime & Law*, *13*, 125-138.

Raynor, P., Kynch, J., Roberts, R., & Merrington, S. (2000). *Risk and need assessment in probation services: An evaluation* (Home Office Research Study 211). London, England: Home Office.

Redondo, S., Sánchez-Meca, J., & Garrido, V. (2002). Crime treatment in Europe: A review of outcome studies. In J. McGuire (Ed.), *Offender rehabilitation and treatment: Effective programmes and policies to reduce re-offending* (pp. 113-141). Chichester, UK: Wiley.

Ross, R., Fabiano, E., & Ewles, C. (1988). Reasoning and rehabilitation. *International Journal for Offender Therapy and Comparative Criminology*, *32*, 29-35.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J. E., Reuter, P., & Bushway, S. D. (1997). *Preventing crime: What works, what doesn't, what's promising*. Washington, DC: Department of Justice, National Institute of Justice.

Singleton, N., Meltzer, H., Gatward, R., Coid, J., & Deasy, D. (1998). *Psychiatric morbidity among prisoners*. London, England: Office of National Statistics.

Stephens, K., Coombs, J., & Debidin, M. (2004). *Black and Asian offenders pathfinder: Implementation report* (Development and Practice Report 24). London, England: Home Office.

Stouthamer-Loeber, M., & Loeber, R. (2002). Lost opportunities for intervention: Undetected markers for the development of serious juvenile delinquency. *Criminal Behaviour and Mental Health*, *12*, 69-82.

Veira, T., Skilling, T., & Peterson-Badali, M. (2009). Matching court-ordered services with treatment needs. *Criminal Justice and Behavior*, *36*, 385-401.

Ward, T., Day, A., Howells, K., & Birgden, A. (2004). The multifactor offender readiness model. *Aggression and Violent Behavior*, *9*, 645-673.

Ward, T., & Stewart, C. (2003). Criminogenic needs and human needs: A theoretical critique. *Psychology, Crime & Law*, *9*, 125-143.

Ward, T., Yates, P., & Willis, G. (2012). The Good Lives Model and the risk, need, responsivity model: A critical response to Andrews, Bonta and Wormith 2011. *Criminal Justice and Behavior*, *39*, 94-112.

Welsh, B. C., & Farrington, D. P. (2011). The benefits and costs of early prevention compared with imprisonment: Toward evidence-based policy. *The Prison Journal*, *91*(3S1), 120-137.

Whitehead, J. T., & Lab, S. P. (1989). A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime & Delinquency*, *26*, 276-295.

Wormith, J. S., Althouse, R., Simpson, M., Reitzel, L. R., Fagan, T. J., & Morgan, R. D. (2007). The rehabilitation and reintegration of offenders: The current landscape and some future directions for correctional criminology. *Criminal Justice and Behavior*, *34*, 880-895.

Wormith, J. S., & Olver, M. (2002). Offender treatment attrition and its relationship with risk, responsivity and recidivism. *Criminal Justice and Behavior*, *29*, 447-471.