



# An Evaluation of Parent Training Interventions in Scotland: The Psychology of Parenting Project (PoPP)

Rob Saunders<sup>1</sup>  · Marita Brack<sup>2</sup> · Brenda Renz<sup>2</sup> · Judy Thomson<sup>2</sup> · Stephen Pilling<sup>1</sup>

Accepted: 29 August 2020  
© The Author(s) 2020

## Abstract

Early-onset behavioural difficulties persisting into the pre-school years, can make young children vulnerable to poor long-term outcomes including the development of conduct disorders, which are linked to significantly higher societal costs. Several parenting interventions have been shown to reduce behavioural difficulties in children and this evaluation presents outcomes from the Psychology of Parenting Project (PoPP), a national implementation programme delivered in early years services in Scotland. This evaluation of service implementation reports on a large cohort of children (2204, age: 2–5 years) whose parents/caregivers participated in PoPP group-based parenting interventions. We explored change in parent-reported Strengths and Difficulties Questionnaire (SDQ) scores following either the Incredible Years Pre-school Basic or the Level 4 Group Triple P interventions. Latent profile analysis (LPA) was used to identify statistically distinct sub-groups of children based on SDQ subscale scores (emotional, conduct, hyperactivity, peer problems and prosocial). Pre- and post-intervention SDQ scores were available for 58% of children. Large intervention effects were reported and analyses showed that 60% of “at-risk” children were no longer scoring in the at-risk range post-interventions. LPA identified four statistically-distinct profiles of children. Children from “low” and “moderate” behavioural problem profiles benefited more from Triple P, whereas “severe” and “hyperactivity-focused” problem profiles displayed better outcomes following Incredible Years. When delivered through a robust implementation scheme, these parenting interventions can be effective in routine service settings and produce clinically important improvements. These findings and the identification of distinct profiles of children who may respond differentially to interventions could guide the planning of future dissemination schemes.

**Keywords** Parent training · Behavioural difficulties · Latent profile analysis · Prevention · Parenting programme

## Highlights

- We examine the Psychology of Parenting Project (PoPP)—a national roll out of evidence-based parenting interventions in Scotland.
- We calculated intervention effects using data from a sample of 2,264 children enrolled in PoPP interventions in the evaluation period.
- 60% of children scoring in the high-risk range of the Strengths and Difficulties Questionnaire pre-assessment were no longer scored in the clinical range post-intervention.
- Overall, we observed similar outcomes when comparing the Incredible Years and Triple P intervention groups.
- These results suggest that PoPP interventions may offer considerable long-term savings when considering potential long-term costs associated with conduct disorders.

---

**Supplementary information** The online version of this article (<https://doi.org/10.1007/s10826-020-01817-y>) contains supplementary material, which is available to authorised users.

---

✉ Rob Saunders  
r.saunders@ucl.ac.uk

Psychology, University College London, 1-19 Torrington Place,  
London, UK

<sup>1</sup> Centre for Outcomes Research and Effectiveness (CORE),  
Research Department of Clinical, Educational and Health

<sup>2</sup> NHS Education for Scotland, 2 Central Quay, 89 Hydepark Street,  
Glasgow, UK

Persistent behavioural difficulties (which can include oppositional behaviours, attention difficulties, hyperactivity and conduct behaviours) in young children are associated with a range of negative life outcomes including school failure, unemployment, substance misuse and mental health problems, and can lead to the onset of the more serious conduct disorder and subsequent poor adjustment in adulthood (Fergusson et al. 2005; Knapp et al. 2011; Patterson et al. 2000). Conduct disorders are characterised by repetitive aggressive, antisocial or problem behaviours that contribute to persistent violations of social and family expectations (National Institute for Health and Care Excellence 2013). Health, social care, education and criminal justice costs are estimated to be ten times higher for children with conduct disorder compared to those without (Parsonage et al. 2014; Sainsbury Centre for Mental Health 2009; Scott et al. 2001).

The “Growing up in Scotland” study (Bradshaw and Tipping 2010) surveyed the level of social, emotional and behavioural difficulties on entry to primary school and found that between 10 and 27% of children in Scotland displayed difficulties above the normal range, and between 5 and 12% of children were considered at “high-risk” for poor outcomes in the future. Parenting interventions aim to improve parenting skills, and address parenting practices that may be contributing to behavioural difficulties (Scott 2008). There is strong evidence from systematic reviews for their effectiveness in reducing reported behavioural difficulties in young children (Epstein et al. 2015; Sanders et al. 2014), especially those aiming to increase parent–child interactions and emotional communication skills (Wyatt Kaminski et al. 2008). Multiple randomised controlled trials of parenting interventions compared to waitlist controls or minimal parenting interventions have found moderate effect sizes (ranging from 0.3 to 0.52) in favour of parenting interventions on measures of a range of behavioural difficulties (Furlong et al. 2012; Hutchings et al. 2007; Scott et al. 2010). This approach represents an opportunity for early intervention, and guidelines for the prevention and management of conduct disorder recommend the use of group-based parenting interventions as an effective method of managing children’s behavioural difficulties (National Institute for Health and Care Excellence 2013).

Treatment gains achieved through the delivery of these interventions have been shown to persist for several years (Drugli et al. 2010). This has further contributed to the cost-effectiveness of these interventions (Edwards et al. 2016) and in the wide-spread dissemination of parent training interventions in England, for example through the Children and Young People’s Increasing Access to Psychological Therapies (CYP-IAPT) initiative and the Parenting Early Intervention Programme (PEIP; (Lindsay and Strand 2013)), where moderate pre–post intervention effects were observed in reducing behavioural difficulties.

NHS Education for Scotland (NES) initiated the Psychology of Parenting Project (PoPP) to increase the availability of evidence-based parent training interventions for Scottish families. In its first 3 years, the primary aim of PoPP was to improve outcomes for 3 and 4-year old children displaying elevated behavioural difficulties (but not more serious conduct disorder). It sought to do this by equipping the existing multi-agency child and family care workforce with the skills to deliver evidence-based parenting interventions, following an identified gap between the number of practitioners who had received core training in evidence-based parenting interventions and the number of groups actually delivered. In response the PoPP implementation framework that was developed drew upon the research and practice within the field of implementation science, and was structured around the three principal implementation drivers proposed by Fixsen et al. (2005): staff competency, organisational supports and leadership. This plan recognised that, in addition to building a confident and competent workforce to deliver the interventions, attention must also be given to the need to create organisational systems and supports to allow staff to deliver the programmes with fidelity, supported by leadership at all levels of the system that was able to be both technical and adaptive. The dissemination plan outlined the educational infrastructure required to complement a re-alignment of some of the considerable resources currently devoted to parenting work across the country, in order to address the current paucity of evidence-based parenting interventions being delivered to families in Scotland.

Between August 2013 and October 2016, the PoPP model had been adopted by 14 Community Planning Partnerships (CPPs) in Scotland, and over 400 multi-agency early years practitioners had completed the 3-day authorised training associated with PoPP-delivered parenting interventions, as well as an enhanced set of training activities developed by the PoPP team. These included training sessions on the PoPP implementation, strength-based communication skills for working with families and the distinct supervision models that were utilised in each of the parenting interventions.

The prevention protocol of the Incredible Years Pre-School intervention (Webster-Stratton 1998) and the Level 4 Group Triple P-Positive Parenting intervention (Sanders 2012) were selected, based on their substantial evidence-base (Furlong et al. 2012; Sanders et al. 2014; Scott et al. 2010), to be delivered by practitioners already in the early years workforce but with newly-dedicated time for this work. A number of potential child moderators have been explored within studies of each intervention. Incredible Years has been found to be effective irrespective of gender, age, and co-occurring attentional issues, as well as levels of anxiety and depression symptoms (Seabra-Santos

et al. 2016; Webster-Stratton 2016). A major systematic review of Triple P interventions identified younger age and higher levels of initial severity associated with greater improvements in child outcomes (Sanders et al. 2014).

The creation of a PoPP database has supported the collection of parent-reported outcomes following parenting interventions and provides an opportunity to evaluate this large-scale roll out of parenting interventions in everyday service settings. It also allows an exploration of potential differences in outcome between two interventions for different groups of children. Previous analysis (Bradshaw and Tipping 2010) has suggested that clusters of children exist based on behavioural symptoms. These authors identified five clusters of children attending primary school, specifically a group that displayed little or no emotional and behavioural issues (the largest group), a group with moderate hyperactivity issues only, a group with particularly high hyperactivity scores only, a group displaying average levels of scores across subscales and a group scoring relatively high across domains of conduct, emotional problems, hyperactivity and peer problems. However, clusters have not been explored in a population of children identified as at-risk, and outcomes have not been explored between these clusters, especially in relation to different interventions. The identification of sub-groups of children with differential outcomes following Incredible Years or Triple P may help to refine the provision of parenting interventions.

The aim of this evaluation of service implementation is to assess the effectiveness of the PoPP-supported interventions in reducing behavioural difficulties in pre-school children in Scotland. In addition, latent profile methods were used to explore the hypothesis that there were different sub-groups of children served by the project, and that these different profiles have differential response to PoPP delivered interventions. This analysis was considered exploratory in approach, but considering previous findings (Bradshaw and Tipping 2010) we hypothesised the existence of a sub-group with difficulties on a number of domains, a group with relatively lower difficulties as well as the potential for a group where issues around hyperactivity were the main concern.

## Method

### Participants

The dataset used for this analysis comes from assessment information for  $n = 2264$  children whose parents enrolled in PoPP-supported parenting interventions between August 2013 and October 2016. Families were referred to a local PoPP team, who would contact the family to discuss the group and establish their interest in taking part. Although it

was recommended that children score over 17 on the Strengths and Difficulties Questionnaire (SDQ; Goodman 2001) (see “Measures” section for further detail) for inclusion, this criteria was not enforced, and the only other inclusion criteria was that children were around 3 to 4 years of age (but families of 2 and 5 years olds could be included if it was considered appropriate). Data was collected prior to the groups starting, at which time they consented to it being used for analysis and evaluation of the PoPP. Pre-intervention assessment data was available for  $n = 2204$  children, the included sample for this analysis.

### Interventions

The Incredible Years Pre-school BASIC intervention (Webster-Stratton 1998) involves the delivery of fourteen face to face, 2-hour long weekly sessions to groups of up to twelve parents. Two practitioners lead discussions of video vignettes, problem-solving exercises and role play activities that address parents’ personal goals. The intervention focuses on strengthening parent–child interactions, nurturing relationships with children, effective management of behaviour, and promoting children’s social, emotional and language development. The Incredible Year manual can be purchased from the Incredible Years website (<http://www.incredibleyears.com>). Level-4 Triple P groups are delivered to a similar number of families with a total of eight sessions delivered, five of which involve 2-hour long face-to-face sessions, delivered by two practitioners, and three 20–30 minute long sessions which are delivered via individual telephone call consultation between one of the practitioners and the parent. During Level-4 Group Triple P (Sanders 2012), parents learn strategies for improving their relationship with their child, increasing the child’s competencies and discouraging unwanted child behaviour. Learning is supported through role play activities, group discussions of video vignettes, and problem-solving exercises focused on the parents’ personal goals. The Triple P manual is available to trained practitioners, with more information available at the Triple P website ([www.triplep.net](http://www.triplep.net)). Fidelity of individual groups was not assessed by the PoPP, but across both interventions fidelity was supported by the use of standardised training, manualised and available materials as well as post-training supervision, accreditation and practitioner networks. Children did not attend either intervention groups.

With these considerations in mind, workforce capacity was created in all but one participating Community Planning Partnership (CPP) for two-thirds of local delivery to involve Triple P groups and one-third to involve Incredible Years groups. This split was made for both pragmatic and financial reasons as Triple P, with 8 sessions, could be delivered four times a year compared to the 14-week

Incredible Years groups. However, it was found that the overall split was closer to one-third Triple P and two-thirds Incredible Years. One CPP made the strategic decision to invest in only one parenting intervention and therefore did not complete any PoPP parenting interventions in the study period. Although the PoPP model was designed so that there was a one-third Incredible Years and two-thirds Triple P split in the delivery of interventions to families, decisions relating to which one of the two interventions would be best suited to a particular family's needs were to be made by the local implementation team. Families were not randomised to interventions and instead practitioners made pragmatic decisions about the allocation of families to interventions depending on a range of factors including the availability of practitioners trained in a specific intervention. In some localities the different interventions were offered on different days and therefore the families were given a choice as to which intervention they attended.

## Measures

We asked all parents to complete a paper copy of the SDQ at both initial assessment and at the final session of the group. A total SDQ score of 17 or more is considered to be in the “high-risk” range of behavioural difficulties. The SDQ has demonstrated good reliability, including for younger children (Mieloo et al. 2012; Stone et al. 2015), but individual items were not available in the current dataset to allow study specific reliability to be calculated. The SDQ is also made up of five sub-scales measuring specific domains of behavioural difficulties which were available to the research team: emotional, conduct problems, hyperactivity, peer problems and prosocial subscales, although the prosocial subscale does not contribute to the total score. The PoPP database also includes basic demographic information on age and gender of children, as well as information about the PoPP delivered parenting interventions.

## Analysis

This analysis included data collected during the PoPP up until October 2016. Post-intervention SDQ scores were not available for  $n = 935$  (42%) of children with initial SDQ information. This was due either to facilitators not gathering and submitting this information or to families not attending the final intervention session. We split the analysis into two samples. A completer sample included all children whose parents attended the last group session and provided post-intervention SDQ scores. A “full-cohort” sample comprised all children with pre-intervention scores only. Where post-intervention SDQ scores were missing in the full-cohort, we carried forward the pre-intervention SDQ score to provide a post-intervention score, making baseline and endpoint

information the same for the  $n = 935$  children with missing outcome data. As no sessional SDQ scores were available, carrying forward initial scores was the only option for this analysis.

We considered three primary outcomes of interest in this evaluation, which were provided for the full sample as well as compared between the intervention types:

- (1) The mean change between pre–post intervention SDQ scores.
- (2) The number of children achieving reliable change in SDQ scores, defined as a change in 7 or more points on the SDQ (Law and Wolpert 2014).
- (3) The number of children “moving out of clinical range”, which is defined as being in the “high-risk” range pre-assessment on the SDQ (score  $\geq 17$ ) then scoring below 17 post-intervention.

A second level of enquiry involved the identification of statistically distinct sub-groups of children whose parents enrolled in PoPP-supported groups. We used latent profile analysis (LPA), an extension of latent class analysis (Hagenaars and McCutcheon 2002; Lazarsfeld and Henry 1968) for this purpose. Although previous analysis of SDQ clustering (Bradshaw and Tipping 2010) used K-means clustering to identify sub-groups based on SDQ assessment scores, LPA methods have shown better performance and are preferred due to the inclusion of model fit statistics (Magidson and Vermunt 2002; Schreiber and Pekarik 2014).

To identify the best fitting model for the data, we compared the Vuong-Lo-Medell-Rubin Likelihood Ratio test (VLMR-LRT; (Lo et al. 2001)) between models, and this was considered alongside the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and entropy values (Asparouhov and Muthén 2012; Geiser 2013). The VLMR-LRT compares the K model (current model, with K number of profiles) to a model with one less profile ( $K - 1$  model), with a significant  $p$  value ( $p < 0.05$ ) indicating better fit for the K model. A  $p$  value  $\geq 0.05$  would suggest the  $K - 1$  model is a better fit, and the more parsimonious model would be preferred. Lower AIC and BIC values between models suggest better fit, and higher entropy values indicate higher classification accuracy for the model. Analysis was conducted in Mplus version-7 (Muthén and Muthén 2012). We then explored outcomes between the identified profiles, with analysis performed in STATA 14 (StataCorp 2015).

## Results

A total of 357 PoPP-supported groups were delivered in the evaluation time frame (218 Incredible Years groups and 139

Triple P groups). The average number of families in Incredible Years groups was 7.2 (standard deviation = 2.59) compared to 5.00 (SD = 1.87) in Triple P groups. The average number of sessions attended for Incredible Years groups was 11.37 (SD = 3.93) and 6.12 (SD = 2.27) for Triple P groups. We present the baseline demographics of children from the completer and full-cohort samples in Table 1. The completer sample consisted of  $n = 1269$  children with post-intervention SDQ information available (57.58% of children with initial SDQ assessment). The families of most children without post-intervention data were reported to have left the intervention groups before the final session (80.86%). The proportion of families leaving the interventions before the final session, and therefore not providing final SDQ scores, was not statistically different between groups (Incredible Years = 36%; Triple P = 32%; odds ratio = 0.843,  $p = 0.082$ ). Over 65% of the child

sample were male, with a mean age of three and a half years. The average total SDQ score pre-intervention was 19.55 (SD = 6.06) for the full-cohort sample, and 19.2 (SD = 6.11) for the completer sample. Mean SDQ subscale scores ranged from 3.1 to 7.1.

A comparison of the descriptive statistics for the full-cohort (pre-intervention SDQ only) and completer samples (pre- and post-intervention scores) suggested that means and proportions were very similar. However, comparison statistics indicated a number of statistical differences between samples. Independent samples  $t$ -tests showed that the mean age of children ( $t = 2.681$ ,  $p = 0.007$ ) and total SDQ at assessment were ( $t = 3.178$ ,  $p = 0.002$ ) significantly higher in the full-cohort sample, as well as for the emotional ( $t = 2.167$ ,  $p = 0.03$ ), conduct ( $t = 3.481$ ,  $p = 0.001$ ) and hyperactivity ( $t = 1.974$ ,  $p = 0.049$ ) sub-scales.

We also performed comparison statistics between children whose parents attended either Incredible Years or Triple P groups to explore where there were differences between frequencies or averages of demographic variables, within both the full-cohort and completer samples. This analysis is presented in Supplementary Table S1. The mean age of children was significantly higher in Triple P groups than Incredible Years, in both the full-cohort (3.52 vs 3.64,  $t = -2.990$ ,  $p = 0.003$ ) and completer samples (3.47 vs 3.61,  $t = -2.479$ ,  $p = 0.013$ ). We found no other significant differences in child characteristics between interventions across either sample ( $p > 0.05$ ).

We present the mean change in SDQ scores between the start and end of the parenting interventions in Table 2. The first section presents change for the full-cohort sample for all children, and by intervention. Results suggest an average reduction of 3.45 points on the SDQ following interventions, with a slightly greater effect found in favour of Incredible Years groups. When only intervention completers were included, mean change was 5.99 points on the SDQ, with large effect sizes indicated for both Incredible Years ( $d = 0.941$ ) and Triple P ( $d = 0.915$ ) groups.

The number of children from both the full-cohort and completer samples who displayed reliable change in SDQ scores (drop of 7 or more points), or those who moved out of clinical range following intervention are presented in

**Table 1** Descriptive statistics of included children at entry to the interventions

Characteristic	Full-cohort sample: $N$ (%) or mean (SD)	Completer sample: $N$ (%) or mean (SD)	$p$ value*
$N$	2204	1269	
Sex			
Male	1443 (65.47)	839 (66.12)	0.364
Female	758 (34.39)	429 (33.81)	
Other	3 (0.14)	1 (0.08)	
Age of child	3.56 (0.86)	3.51 (0.86)	0.007
SDQ scores			
Total score	19.55 (6.06)	19.2 (6.11)	0.002
Subscale: emotional	3.39 (2.42)	3.29 (2.4)	0.030
Subscale: conduct	5.87 (2.36)	5.71 (2.37)	0.001
Subscale: hyperactivity	7.1 (2.44)	7.01 (2.43)	0.049
Subscale: peer problems	3.15 (2.08)	3.1 (2.05)	0.227
Subscale: prosocial	5.89 (2.8)	5.88 (2.25)	0.828

\* $p$  values from chi-square tests and independent samples  $t$ -tests

**Table 2** Mean pre, post and change in SDQ scores

Intervention (sample)	$n$	Pre-score Mean (SD)	Post-score Mean (SD)	Change score Mean (SD)	Effect size ( $d$ )
Full-cohort—all children	2204	19.55 (6.06)	16.1 (7.24)	3.45 (5.63)	0.517
Full-cohort—Incredible Years	1523	19.64 (6.17)	16.06 (7.37)	3.58 (5.79)	0.527
Full-cohort—Triple P	681	19.34 (5.81)	16.17 (6.92)	3.16 (5.25)	0.496
Completers—all children	1269	19.2 (6.11)	13.21 (6.71)	5.99 (6.31)	0.933
Completers—Incredible Years	890	19.27 (6.2)	13.15 (6.8)	6.12 (6.47)	0.941
Completers—Triple P	379	19.03 (5.9)	13.35 (6.5)	5.68 (5.94)	0.915



**Table 3** Number and percentage of children achieving reliable change and moving out of the “high-risk” range

Intervention	Reliable change		Moving out of high-risk range	
	Total children in analysis	% reliable change	Total children in analysis	% moving out of high-risk range
Full-cohort sample				
All children	2204	25.45%	1582	31.80%
Incredible Years	1523	26.59%	1097	32.82%
Triple P	681	22.91%	485	29.48%
Completer sample				
All children	1269	44.21%	839	59.95%
Incredible Years	890	45.51%	591	60.91%
Triple P	379	41.16%	248	57.66%

**Table 4** Number and percentage of children achieving reliable change and moving out of the “high-risk” range

Number of profiles	Log-likelihood	AIC	BIC	Adj-BIC	VLMR-LRT <i>p</i> value	Entropy	Classification (% per profile)
1-Profile	−23838	47695	47752	47720	n/a	n/a	n/a
2-Profile	−23328	46688	46779	46728	<0.05	0.694	34/66
3-Profile	−23202	46449	46573	46503	<0.05	0.601	26/28/46
<b>4-Profile</b>	<b>−23147</b>	<b>46351</b>	<b>46509</b>	<b>46420</b>	<b>&lt;0.05</b>	<b>0.634</b>	<b>13/18/40/29</b>
5-Profile	−23091	46249	46441	46333	0.053	0.663	11/24/8/36/21
6-Profile	−23042	46164	46391	46264	0.343	0.649	17/23/8/29/17/6

Selected model are indicated in bold

Table 3. It was found that 44% of children in the completer sample, and 25% of the full-cohort sample (where pre-intervention SDQ scores were carried forward for children with missing post-intervention data) achieved reliable change. A total of  $n = 1582$  from the full-cohort and  $n = 839$  children in the completer sample scored in the high-risk range of the SDQ at pre-intervention assessment and therefore could be included in the analysis. We observed that moving out of clinical range was reported for approaching 60% of completers, and 32% of the full-cohort who were previously in the “high-risk” range. The proportion of positive outcomes was slightly higher following Incredible Years groups.

LPA was then performed on the SDQ sub-scale dataset to identify statistically-distinct profiles of children served through the PoPP scheme. Of the initial  $n = 2204$  children,  $n = 90$  had no SDQ subscale scores available, and therefore the included sample consisted of  $n = 2114$  children with pre-intervention SDQ subscale scores. Model fit statistics are displayed in Table 4. The VLMR-LRT indicated statistically significant  $p$  values for the 2-profile, 3-profile and 4-profile solutions, before the 5-profile solution resulted in a  $p$  value  $> 0.05$ , suggesting that the 4-profile solution was the best fit for the data. The AIC and BIC values decreased from the 2- to 6-profile solutions, with the rate decelerating. The VLMR-LRT  $p$  value for the 5-profile solution was only just non-significant ( $p = 0.053$ ) and therefore the 4- and

5-profile solutions were compared. The fifth profile (in the 5-profile solution) was found to have average subscale scores in between two other profiles, providing little additional clinical information, and therefore the 4-class solution was considered optimal given the non-significant  $p$  value for the VLMR-LRT of the 5-profile solution (Nylund et al. 2007). Included children were allocated to the latent profile (LP) to which they had the highest probability of membership.

The distribution of children across LPs and mean subscale scores for each LP are displayed in Supplementary Table S2. The four profiles are described below and are also presented graphically in Supplementary Fig. S3.

- (1) LP1 (low problem behaviours)—children in this profile score relatively low on all subscales of the SDQ and could be considered a group of children with low problem behaviours.
- (2) LP2 (moderate problem behaviours)—members of LP2 show moderate levels of problem behaviour. Making up 12.7% of children, this profile includes the smallest number of participants.
- (3) LP3 (hyperactivity problems)—LP3 is distinguished by its considerably higher mean hyperactivity subscale scores when compared to all other subscales. The mean subscale scores on all other domains are similar to those of LP2 (moderate), and therefore this profile is considered a hyperactivity problem specific

**Table 5** Moving out of high-risk range by latent profile and intervention

Latent profile	All Children		Incredible Years		Triple P	
	Total children	% out of high-risk range	Total children	% out of high-risk range	Total children	% out of high-risk range
Low	17	76.47%	12	75.00%	5	80.00%
Moderate	97	70.10%	64	67.19%	33	75.76%
Hyperactivity	355	71.83%	266	72.93%	89	68.54%
Severe	335	42.69%	244	44.67%	91	37.36%
Total	804	59.58%	586	60.58%	218	56.88%

subgroup.

- (4) LP4 (severe problem behaviours)—children in this profile display the most severe problem behaviours compared to the other profiles. All mean subscale scores except emotional symptoms are in the high-risk range, but hyperactivity and conduct are particularly severe.

In the next stage, behavioural outcomes were explored between the profiles, as well as in response to the two different intervention groups, using only children with post-intervention data available (the completer sample). Table 5 presents the proportion of children from each profile who were no longer scoring in the high-risk range following intervention. The table indicates that the number of children who were scoring high-risk initially was limited for the low and moderate profiles, which is due to the lower total SDQ scores for these profiles of children.

Overall the percentage of children moving out of the clinical range was more than 70% for the low, moderate and hyperactivity profiles, whilst it was 43% for children from the severe profile of behavioural difficulties. Comparing outcomes between Incredible Years and Triple P groups suggests that for the low and moderate profiles, Triple P groups show more positive outcomes, whereas slightly more benefit was indicated for hyperactivity and severe profile children following Incredible Years groups, although differences were not statistically significant ( $p > 0.05$ ). It should be noted, however, that the low numbers of children included in this analysis for the low and moderate profile groups means that results should be interpreted with caution.

We then explored the mean change in SDQ total between profiles, and present results in the left-hand columns of Table 6. The findings suggest that effect sizes were highest in the severe profile, followed by the hyperactivity profile, then the moderate and low profiles, which was consistent across interventions. These larger effect sizes in the profiles with higher severity at baseline may be in part due to regression to the mean, as these children would have more available change in scale scores than children with lower scores. Effect sizes for Incredible Years groups were slightly higher than for Triple P groups, although

differences in mean change scores were not statistically significant ( $p > 0.05$ ). The smaller effect size for the low behavioural difficulties profile is likely due to the lower initial assessment symptom scores resulting in less available change in symptoms.

As the conduct and hyperactivity sub-scales were the highest scoring at initial assessment, and contributed to the distribution of the profiles, change in these sub-scales was also explored in this analysis and presented in the right-hand columns of Table 6. The results suggest that Triple P interventions produced more conduct subscale change for the moderate profile, whereas effect sizes are larger for the severe profile when Incredible Years were delivered. The hyperactivity subscale change showed a larger effect size for the Low profile when Triple P interventions were received, but more change was indicated for the hyperactivity profile following Incredible Years interventions on this SDQ subscale. Although differences were observed, these were not statistically significant.

## Discussion

The results of this evaluation suggest that the evidence-based parenting interventions delivered as part of the PoPP implementation scheme in everyday service settings have had a positive effect in reducing behavioural difficulties in children in Scotland, with effect sizes comparable to controlled trials. We found that nearly 60% of children with post-intervention assessment information who were initially scoring “high-risk” for problem behaviours were no longer high-risk following intervention. A conservative analysis performed on the full-cohort sample, whereby children who did not provide post-intervention data were assumed to still be high-risk, resulted in a drop from 60% of completers to 32% of the full-cohort sample moving out of clinical range (33% Incredible Years, 29% Triple P). Considering the potential long-term costs associated with conduct disorders (Scott et al. 2001), the effectiveness of the Incredible Years and Triple P groups implemented within the PoPP framework suggests the potential for considerable long-term savings.

**Table 6** Mean change in Total SDQ score, conduct and hyperactivity subscale for each latent profile by intervention type

Latent profile	<i>n</i>	Total SDQ score				Conduct subscale				Hyperactivity subscale			
		Mean SDQ change	SD change	Mean change ( <i>d</i> )	Effect size ( <i>d</i> )	Mean SDQ change	SD change	Mean change ( <i>d</i> )	Effect size ( <i>d</i> )	Mean SDQ change	SD change	Mean change ( <i>d</i> )	Effect size ( <i>d</i> )
All children													
Low	231	2.81	4.61	0.67	0.89	2.01	2.01	0.52	0.71	1.91	0.42		
Moderate	162	5.75	6.04	1.16	2.4	2.3	2.3	1.21	0.63	2.23	0.36		
Hyperactivity	482	6.47	5.79	1.30	1.99	2.19	2.19	1.05	2.69	2.3	1.42		
Severe	341	7.68	7.22	1.35	2.83	2.66	2.66	1.30	2.43	2.6	1.22		
Total	1216	6.02	6.29	0.94	2.07	2.4	2.4	0.89	1.96	2.48	0.78		
Incredible Years													
Low	183	2.85	4.73	0.68	0.91	2.06	2.06	0.52	0.63	2.03	0.36		
Moderate	98	5.82	6.24	1.18	2.23	2.21	2.21	1.11	0.7	2.25	0.41		
Hyperactivity	348	6.62	5.85	1.34	1.97	2.17	2.17	1.05	2.8	2.36	1.47		
Severe	250	7.99	7.5	1.37	2.9	2.71	2.71	1.31	2.51	2.68	1.23		
Total	879	6.14	6.46	0.94	2.04	2.42	2.42	0.86	2.03	2.56	0.79		
Triple P													
Low	48	2.65	4.17	0.66	0.79	1.83	1.83	0.55	1.02	1.36	0.72		
Moderate	64	5.64	5.78	1.13	2.67	2.44	2.44	1.39	0.51	2.2	0.27		
Hyperactivity	134	6.07	5.63	1.20	2.02	2.24	2.24	1.06	2.4	2.14	1.28		
Severe	91	6.85	6.34	1.30	2.64	2.49	2.49	1.27	2.2	2.38	1.22		
Total	337	5.71	5.81	0.93	2.13	2.37	2.37	0.97	1.78	2.25	0.75		



The effect sizes for pre–post change in total SDQ scores ranged from 0.915 to 0.941 for intervention completers and 0.496 to 0.527 for the full-cohort. These are equivalent, if not larger, than effect sizes indicated in other research-oriented implementation evaluations and controlled trials of parenting groups (Lindsay and Strand 2013; Sanders et al. 2014; Scott et al. 2010). This may, in part, be due to the younger age of children included in the PoPP, as this has been associated with better outcomes (Sanders et al. 2014). Equally, this strong performance may be associated with the robustness of the PoPP implementation scheme, the aim of which was to support practitioners to provide interventions that were consistently delivered to a high standard, even in routine settings. Anecdotally parents spoke positively about changes in their child’s behaviour, their relationship with their child, and their confidence and skills as parents. PoPP practitioners reported seeing tremendous changes in families, but also in their own skills and practice. Feedback evaluations on PoPP delivered groups were provided by 588 families, with 99.66% reporting they would recommend the groups received to a friend, 96.77% agreeing that the groups improved their relationship with their child and 97.69% reporting their family life has benefitted. In response to the question “how have you personally benefitted from the groups?”, quotes included:

“My relationship with my daughter has become more happier & a stronger bond”

“I feel more calm & in control, the general household is much calmer & happier”

“more confidence in myself & my child”.

Comparison of the parenting intervention types showed that overall outcomes were similar, with a slight advantage of the Incredible Years interventions compared to Triple P. It is not possible to tell from this dataset what factors may be contributing to the reported differences. They could, for example be due to the Incredible Years intervention being provided over 14-weeks compared to the 8-week Triple P groups, to programme-specific factors or to local implementation processes.

We found that 35% of children whose parents attended at least the first group session did not have a final SDQ score recorded, with little difference between Incredible Years and Triple P interventions, despite their different attendance demands. This suggests a higher than usual non-completion rate for PoPP-supported groups than that reported in controlled trials of parenting interventions which have indicated between 18 and 25% dropout (Scott

et al. 2001; Scott et al. 2010). It is, however, still considerably lower than non-completion rates reported in other national evaluations of parenting interventions such as the PEIP (46%, (Lindsay and Strand 2013)). The higher non-completion in PoPP-supported groups compared to controlled trials is likely due to the more naturalistic settings in which the PoPP-supported groups were delivered. It may also be confounded by a lack of data to clarify the proportion of parents who did attend the final group session but did not provide a final SDQ score for their child as opposed to those who had stopped attending the group. Nevertheless, as the reduction between outcomes between the full-cohort and completers samples was significant, methods of reducing the proportion of families leaving interventions early may further improve PoPP outcomes.

Using SDQ subscale scores, LPA identified four statistically-distinct profiles of children whose parents attended PoPP-supported groups (low, moderate, severe and hyperactivity problems profiles). The profiles have some broad similarities to profiles of SDQ subscales identified in previous analyses (Bradshaw and Tipping 2010), in that the current study also identified a low, moderate and high group, as well as the existence of a group with noticeable higher hyperactivity subscale scores. The main differences were that a fifth profile, identified in previous work as sub-group with above average hyperactivity and that might be considered boisterous without having specific difficulties. This may be due to the previous research surveying all primary school children rather than a group identified as having at risk behavioural difficulties as in the current analysis. This may also explain why the low severity group were the biggest group in the previous study but not for the current study, and why the mean subscales of the “low” group in the current analysis were much higher than those in the previously identified cluster.

We observed differential outcomes between these profiles, and high-risk members of the severe profile group were less likely to move out of the clinical range (43%) compared to the three other profiles, where likelihood was over 70%. It should be noted that these differences were not statistically significant, indicating further research into differential outcomes between profiles is warranted. Large pre–post intervention effect sizes were observed for all profiles except the low problem profile. However, this may be accounted for by the lower initial SDQ score for this profile resulting in limited available change in measured behaviours. Some intervention-specific differential outcomes were also observed for children based on their different SDQ problem profiles. Local services could use these findings to support the implementation and sustainability of this form of effective early intervention; to take decisions about how they apportion their delivery capacity for each

intervention in the future; and to guide a more tailored delivery of these parenting interventions. As only about half of the children with elevated levels of behavioural difficulty at this age would be expected to develop conduct disorder, efficient use of resources demands as accurate prediction as possible of those children who are most at risk and who are most likely to benefit.

This naturalistic evaluation of parent training interventions had a number of limitations which need to be considered in relation to the findings discussed in this study. Firstly, no measures of staff adherence in the parenting groups was included and it is possible that there were local differences in delivery across PoPP sites. However, all PoPP practitioners received the same levels of training, and each site benefited from equivalent and standardised implementation support measures. This included additional clinical skills consultation days, supervision structures, as well as the implementation support given to each CPPs by the PoPP implementation team (via regular email and telephone consultation and review meetings). A second limitation is that only pre- and post-intervention SDQ data was available, with no sessional data collection. Collecting assessment information at every session, or even midway through would have allowed further exploration of change in scores through interventions. Instead, the conservative analysis performed using the full-cohort sample presumes that all children without post-intervention SDQ scores did not show any benefit. Only SDQ subscale scores were available, meaning the reliability of scale could not be assessed in the sample. Lastly, the number of baseline characteristics of children and families was limited and collecting additional information would provide the opportunity to further explore characteristics of children who were more likely to benefit from one intervention as opposed to the other.

## Conclusion

Overall, the parent training interventions delivered through the PoPP implementation scheme were associated with significant reductions in behavioural difficulties in young children in various localities in Scotland. The effects reported are equivalent, and at times larger, than those reported by randomised controlled trials of parenting interventions. The identification of profiles of children based on SDQ subscale scores could also be used to inform future service delivery decisions. The naturalistic settings of this evaluation have associated limitations such as a lack of ability to assess practitioner fidelity, but provide support for the use of parent training programme for reducing behavioural difficulties in at risk pre-school children. Further

research is needed to explore the utility of the profiles identified in supporting intervention allocation.

**Acknowledgements** We would like to thank all the families who participated in the PoPP groups that are discussed in this paper and the practitioners and the wider implementation teams from the 14 contributing Community Planning Partnerships who worked with us to deliver these interventions (West Lothian, Falkirk, Stirling, Clackmannanshire, Fife, Argyll and Bute, Perth and Kinross, City of Edinburgh, Scottish Borders, Highland, East Lothian, North Ayrshire, Aberdeenshire and East Renfrewshire).

**Funding** The Psychology of Parenting Project is funded through the Scottish Government Mental Health Division. S.P. is supported by funding from the UCLH National Institute for Health (NIHR) Biomedical Research Centre.

## Compliance with Ethical Standards

**Conflict of Interest** B.R. is a Mentor within the Incredible Years and delivers training in the Incredible Years intervention. All other authors declare that they have no conflicts of interest.

**Ethical Approval** Ethical standards: as the present study is a service evaluation, the NHS Health Research Authority does not classify this study as research, and as such ethical approval was not required.

**Informed Consent** Informed consent was gained from all parents for the collection of their data to be used for analysis to determine if the groups were effective for families and why they were effective.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Asparouhov, T., & Muthén, B. (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes. *Mplus Web Notes*, 14. <https://www.statmodel.com/examples/webnotes/webnote14.pdf>.
- Bradshaw, P., & Tipping, S. (2010). *Growing Up in Scotland: Children's social, emotional and behavioural characteristics at entry to primary school*. Edinburgh: The Scottish Government.
- Drugli, M. B., Larsson, B., Fossum, S., & Mørch, W. T. (2010). Five- to six-year outcome and its prediction for children with ODD/CD treated with parent training. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 51(5), 559–566. <https://doi.org/10.1111/j.1469-7610.2009.02178.x>.

- Edwards, R. T., Jones, C., Berry, V., Charles, J., Linck, P., Bywater, T., & Hutchings, J. (2016). Incredible Years parenting programme: cost-effectiveness and implementation. *Journal of Children's Services, 11*(1), 54–72. <https://doi.org/10.1108/JCS-02-2015-0005>.
- Epstein, R. A., Fonnesebeck, C., Potter, S., Rizzone, K. H., & McPheeters, M. (2015). Psychosocial interventions for child disruptive behaviors: a meta-analysis. *Pediatrics, 136*(5), 947–960. <https://doi.org/10.1542/peds.2015-2577>.
- Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*(8), 837–849. <https://doi.org/10.1111/j.1469-7610.2004.00387.x>.
- Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: a synthesis of the literature*. Louis de la Parte Florida Mental Health Institute Publication. [https://doi.org/10.1016/S0043-1648\(02\)00098-4](https://doi.org/10.1016/S0043-1648(02)00098-4).
- Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *The Cochrane Database Of Systematic Reviews, 15*(2), CD008225. <https://doi.org/10.1002/ebch.1905>.
- Geiser, C. (2013). *Data analysis with Mplus*. New York, NY: Guilford.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(22), 1337–1345. <https://doi.org/10.1073/pnas.0711961105>.
- Hagenaars, J., & McCutcheon, A. (2002). *Applied latent class analysis*. Cambridge: Cambridge University Press.
- Hutchings, J., Gardner, F., Bywater, T., Daley, D., Whitaker, C., Jones, K., & Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *British Medical Journal, 334*, 678. <https://doi.org/10.1136/bmj.39126.620799.55>.
- Knapp, M., King, D., Healey, A., & Thomas, C. (2011). Economic outcomes in adulthood and their associations with antisocial conduct, attention deficit and anxiety problems in childhood. *Journal of Mental Health Policy and Economics, 14*(3), 137–147.
- Law, D., & Wolpert, M. (2014). *Guide to using outcomes and feedback tools with children, young people and families formally known as COOP document*. London: Child Outcomes Research Consortium. [https://www.corc.uk.net/media/2112/201404guide\\_to\\_using\\_outcomes\\_measures\\_and\\_feedback\\_tools-updated.pdf](https://www.corc.uk.net/media/2112/201404guide_to_using_outcomes_measures_and_feedback_tools-updated.pdf).
- Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.
- Lindsay, G., & Strand, S. (2013). Evaluation of the national roll-out of parenting programmes across England: the parenting early intervention programme (PEIP). *BMC Public Health, 13*(972). <https://doi.org/10.1186/1471-2458-13-972>.
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767–778.
- Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: a comparison with K-means. *Canadian Journal of Marketing Research, 20*(1), 37–44.
- Mieloo, C., Raat, H., van Oort, F., Bevaart, F., Vogel, I., Donker, M., & Jansen, W. (2012). Validity and reliability of the strengths and Difficulties Questionnaire in 5–6 year olds: differences by gender or by parental education? *PLoS ONE, 7*(5). <https://doi.org/10.1371/journal.pone.0036805>.
- Muthén, L., & Muthén, B. (2012). *Mplus version 7 user's guide*. Los Angeles, CA: Muthén & Muthén.
- National Institute for Health and Care Excellence (2013). *Antisocial behavior and conduct disorders in children and young people—the NICE guideline on recognition intervention and management*. National Institute for Health and Care Excellence. <http://www.nice.org.uk/guidance/cg158/chapter/2-research-recommendations#parent-training-programmes-for-children-aged-12-years-and-over-with-a-conduct-disorder>.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Structural Equation Modeling, 14*(4), 535–569. <https://doi.org/10.1080/10705510701575396>.
- Parsonage, M., Khan, L., & Saunders, A. (2014). *Building a better future: the lifetime costs of childhood behaviour problems and the benefits of early intervention*. <https://www.centreformentalhealth.org.uk/sites/default/files/2018-09/buildingabetterfuture.pdf>.
- Patterson, G. R., DeGarmo, D. S., & Knutson, N. (2000). Hyperactive and antisocial behaviors: comorbid or two points in the same process? *Development and Psychopathology, 12*(1), 91–106. <https://doi.org/10.1017/S0954579400001061>.
- Sainsbury Centre for Mental Health (2009). *The chance of a lifetime: preventing early conduct problems and reducing crime*. London: Sainsbury Centre for Mental Health. <http://www.ohrn.nhs.uk/resource/policy/SCMHThechanceofalifetime.pdf>.
- Sanders, M. R. (2012). Development, evaluation, and multinational dissemination of the Triple P-Positive Parenting Program. *Annual Review of Clinical Psychology, 8*, 345–379. <https://doi.org/10.1146/annurev-clinpsy-032511-143104>.
- Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: a systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review, 34*(4), 337–357. <https://doi.org/10.1016/j.cpr.2014.04.003>.
- Schreiber, J. B., & Pekarik, A. J. (2014). Technical note: using latent class analysis versus K-means or hierarchical clustering to understand museum visitors. *Curator: The Museum Journal, 57*(1), 45–59. <https://doi.org/10.1111/cura.12050>.
- Scott, S. (2008). An update on interventions for conduct disorder. *Advances in Psychiatric Treatment, 14*(1), 61–70. <https://doi.org/10.1192/apt.bp.106.002626>.
- Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into. *BMJ, 323*, 191. <https://doi.org/10.1136/bmj.323.7306.191>.
- Scott, S., Spender, Q., Doolan, M., Jacobs, B., & Aspland, H. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *BMJ, 323*, 1–7. <https://doi.org/10.1136/bmj.323.7306.194>.
- Scott, S., Sylva, K., Doolan, M., Price, J., Jacobs, B., Crook, C., & Landau, S. (2010). Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: the SPOKES project. *Journal of Child Psychology and Psychiatry, 51*(1), 48–57. <https://doi.org/10.1111/j.1469-7610.2009.02127.x>.
- Seabra-Santos, M. J., Gaspar, M. F., Azevedo, A. F., Homem, T. C., Guerra, J., Martins, V., & Moura-Ramos, M. (2016). Incredible Years parent training: what changes, for whom, how, for how long? *Journal of Applied Developmental Psychology, 57*. <https://doi.org/10.1016/j.appdev.2016.04.004>.
- StataCorp (2015). Stata statistical software: release 14. <https://doi.org/10.2307/2234838>.
- Stone, L. L., Janssens, J. M. A. M., Vermulst, A. A., Van Der Maten, M., Engels, R. C. M. E., & Otten, R. (2015). The Strengths and Difficulties Questionnaire: psychometric properties of the parent and teacher version in children aged 4–7. *BMC Psychology, 3*(1), 4. <https://doi.org/10.1186/s40359-015-0061-8>.
- Webster-Stratton, C. (1998). Preventing conduct problems in head start children: strengthening parenting competencies. *Journal of Consulting and Clinical Psychology, 66*(5), 715–730. <https://doi.org/10.1037/0022-006X.66.5.715>.

- Webster-Stratton, C. (2016). The Incredible Years® series: a developmental approach. In *Family-centered prevention programs for children and adolescents: theory, research, and large-scale dissemination* (pp. 42–67). [http://www.incredibleyears.com/wp-content/uploads/IY-Series\\_A-Developmental-Approach.pdf](http://www.incredibleyears.com/wp-content/uploads/IY-Series_A-Developmental-Approach.pdf).
- Wyatt Kaminski, J., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*. <https://doi.org/10.1007/s10802-007-9201-9>.